



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification⁵ : A61K 37/00, 35/14, C07H 15/12 C07H 17/00, C07K 3/00, 15/00 C07K 17/00, C12N 1/00, 15/00	A1	(11) International Publication Number: WO 91/04036 (43) International Publication Date: 4 April 1991 (04.04.91)
(21) International Application Number: PCT/US90/05129 (22) International Filing Date: 11 September 1990 (11.09.90) (30) Priority data: 406,535 13 September 1989 (13.09.89) US (71) Applicant: ML TECHNOLOGY VENTURES, L.P. [US/ US]; World Financial Center, 18th Floor, North Tower, New York, NY 10281 (US). (72) Inventors: GABE, Jeffrey ; 2364 Fulton, San Francisco, CA 94118 (US). DRAGON, Elizabeth ; 42 Park Lane, Orinda, CA 94563 (US). McCAMAN, Michael ; 746 Cherry Avenue, San Bruno, CA 94066 (US).		(74) Agents: OLSTEIN, Elliot, M. et al.; Carella, Byrne, Bain, Gilfillan, Cecchi & Stewart, 6 Becker Farm Road, Rose- land, NJ 07068 (US). (81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent)*, DK (European patent), ES (European patent), FR (Eu- ropean patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European pa- tent), SE (European patent). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the</i> <i>claims and to be republished in the event of the receipt of</i> <i>amendments.</i>
(54) Title: TREPONEMA HYODYSENTERIAE ANTIGENS HAVING A MOLECULAR WEIGHT OF 39kDa AND DNA ENCODING THEREFOR (57) Abstract A family of <i>T. hyodysenteriae</i> 39 kDa antigens are produced by recombinant techniques. Seven different genes and anti- gens have been identified. Such antigens may be used in a vaccine.		

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MC	Monaco
AU	Australia	FI	Finland	MG	Madagascar
BB	Barbados	FR	France	ML	Mali
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GR	Greece	NL	Netherlands
BJ	Benin	HU	Hungary	NO	Norway
BR	Brazil	IT	Italy	PL	Poland
CA	Canada	JP	Japan	RO	Romania
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
DE	Germany	LK	Sri Lanka	TD	Chad
DK	Denmark	LU	Luxembourg	TC	Togo
				US	United States of America

TREPONEMA HYODYSENTERIAE ANTIGENS
HAVING A MOLECULAR WEIGHT OF 39kDa
AND DNA ENCODING THEREFOR

This invention relates to *Treponema hyodysenteriae* and more particularly to *Treponema hyodysenteriae* (*T. hyo.*) antigens, genes encoding for such antigens, cells genetically engineered with DNA encoding for such antigens and uses for such antigens. Still more particularly, this invention relates to *Treponema hyodysenteriae* antigens having a molecular weight of 39kDa and to the production thereof by recombinant techniques.

Swine dysentery is a severe, infectious disease found in all major pig-rearing countries. The symptoms of swine dysentery are severe mucohemorrhagic diarrhea, dehydration and weight loss.

The present invention is directed to certain antigens which are useful in determining and/or treating *Treponema hyodysenteriae* and to recombinant or genetic engineering techniques for producing such antigens.

In accordance with one aspect of the present invention, there is provided a protein which is

-2-

capable of eliciting at least one antibody which recognizes an epitope of at least one T.hyo antigen having a molecular weight of about 39kDa.

Applicant has found that T.hyo includes DNA which encodes for a plurality of proteins each having a molecular weight of about 39kDa. Still more particularly, Applicant has found that there are at least eight different genes, each of which encodes for a T.hyo protein having a molecular weight of about 39 kDa. The protein products encoded by such genes have been found to have conserved regions which are interspersed with variable regions. It has been found that the variable regions are generally located in the more hydrophilic portions of the protein whereas the conserved regions are located in the more hydrophobic portions of the protein.

A comparison of the predicted amino acid sequences from the mature peptides encoded by the 39kDa family genes is found in Appendix 1. The peptide sequence corresponding to the signal peptides of these proteins (Appendices 2A and 2B) has been removed for the purposes of this comparison. Examination of this comparison reveals each gene encodes a protein product of similar molecular weight and that there are regions of conserved protein sequence punctuated by regions of variable sequence. The conserved regions are generally in the more hydrophobic portions of the proteins while the variable regions tend to be in the more hydrophilic portions. (Kyte & Doolittle, Journal of Molecular Biology, 157, 105 (1982))

Thus, in accordance with one aspect of the present invention, there is provided eight different antigens (or fragments or analogs thereof), which are T.hyo antigens which have a molecular weight of about

-3-

39 kDa. Such seven genes are hereinafter sometimes referred to as genes 1-8 or copies 1-8.

In accordance with another aspect of the present invention, there is provided at least eight different genes, each of which encodes for a different T.hyo antigen having a molecular weight of about 39 kDa.

In accordance with yet another aspect of the present invention, there is provided an expression or cloning vehicle which includes a DNA sequence which encodes for a T.hyo antigen (or fragment or analog thereof), which has a molecular weight of about 39 kDa.

In accordance with yet a further aspect of the present invention, there is provided a host cell or organism which is genetically engineered with DNA which encodes for a T.hyo antigen (or fragment or derivative thereof), which has a molecular weight of about 39 kDa.

The molecular weight for characterizing the 39 kDa T. hyo. antigen or protein is obtained by discontinuous polyacrylamide gel electrophoresis using the SDS buffer system described by Laemmli, Nature, 227:680-85 (London, 1970) with an acrylamide concentration of 10-17% and a bis-acrylamide to acrylamide ratio of 1:29.

Thus, Applicant has found that there are at least eight different T.hyo antigens, each of which has a molecular weight of about 39 kDa, which are encoded by eight different genes.

The DNA sequence may encode for a protein which is the entire 39 kDa antigen, or a fragment or derivative of the antigen, or a fusion product of the antigen or fragment and another protein, provided that the protein which is produced from such DNA sequence elicits antibodies after immunization which

-4-

recognize an epitope of a 39 kDa T. hyo. antigen. Thus, for example, the DNA sequence may encode for a protein which is or contains within it a fragment of a 39 kDa antigen provided that such fragment generates antibodies which recognize an epitope of a 39 kDa antigen.

Similarly, the DNA sequence may encode for a protein which is a derivative of the antigen e.g., a mutation of one or more amino acids in the peptide chain, as long as such derivative elicits antibodies which recognize an epitope(s) of a 39 kDa T.hyo. antigen as hereinabove described.

The DNA sequence may encode a protein which is a fusion product of (i) a protein which produces antibodies which recognize an epitope(s) of a noted 39 kDa T.hyo. antigen and (ii) another protein (for example chymosin).

The 39 kDa antigens may vary somewhat between specific strains of T. hyo. Thus, for example, the 39 kDa proteins of serotype B204 have minor differences from those of serotype B234; however, such antigens, as well as the genes encoding such antigens are essentially identical to each other.

As a result, the term "DNA sequence which encodes for a protein which produces antibodies which recognize an epitope(s) of a noted 39 kDa T. hyo. antigen" encompasses DNA sequences which encode for and/or express in appropriate transformed cells, proteins which may be the appropriate antigen, antigen fragment, antigen derivative or a fusion product of such antigen, antigen fragment or antigen derivative with another protein.

It is also to be understood that the DNA sequence present in the vector when introduced into a cell may express only a portion of the protein which

-5-

is encoded by such DNA sequence, and such DNA sequence is within the noted terminology, provided that the protein portion expressed elicits antibodies which recognize an epitope(s) of one or more of the noted 39 kDa T. hyo. antigens. For example, the DNA sequence may encode for the entire antigen; however, the expressed protein is a fragment of the antigen.

The term "gene (1, 2, 3, 4, 5, 6, 7 or 8) encoding a T. hyo. 39 kDa protein" means the entire or full length gene sequence or an analog, fragment or derivative thereof which encodes a protein which is capable of eliciting at least one antibody which recognizes at least one epitope of the full length T. hyo. 39 kDa antigen encoded by such full length gene.

The term "protein encoded by gene 1, 2, 3, 4, 5, 6, 7 or 8" means a T. hyo. 39 kDa protein encoded by the entire or full length gene or an analogue, fragment or derivative of such protein which is capable of eliciting at least one antibody which recognizes at least one epitope of the full length T. hyo. 39 kDa antigen encoded by such full length gene.

The term "39 kDa T. hyo. antigen or protein" means a T. hyo. antigen or protein having a molecular weight of about 39 kDa.

The appropriate DNA sequence may be included in any of a wide variety of vectors or plasmids. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences; e.g., derivatives of SV40; bacterial plasmids; phage DNA's; yeast plasmids; vectors derived from combinations of plasmids and phage DNAs, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an

-6-

appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the E. coli. lac or trp, the phage lambda PL promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably contain a gene to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in E. coli.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein. As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as E. coli, Salmonella typhimurium; fungal cells, such as yeast; animal cells such as CHO or Bowes melanoma; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

-7-

As hereinabove indicated, the expression vehicle including the appropriate DNA sequence inserted at the selected site may include a DNA or gene sequence which is not part of the gene coding for the protein which is capable of producing antibodies which recognize an epitope(s) of the noted T. hyo. antigen(s). For example, the desired DNA sequence may be fused in the same reading frame to a DNA sequence which aids in expression or improves purification or permits expression of the appropriate protein.

When seeking to develop a vaccine neutralizing or protective antibodies could be targeted towards discontinuous, conformation-dependent epitopes of the native antigen. One must therefore consider whether the protein obtained from the recombinant expression system might have a three dimensional structure (conformation) which differs substantially from that of the original protein molecule in its natural environment. Thus, dependent on the immunogenic properties of the isolated proteins, one might need to renature it to restore the appropriate molecular conformation. Numerous methods for renaturation of proteins can be found in the scientific literature and include; (1) denaturation (unfolding) of improperly folded proteins using agents such as alkali, chaotropes, organic solvents and ionic detergents followed by a renaturation step achieved by dilution, dialysis, or pH adjustment to remove the denaturant, and (2) reconstitution of proteins into a lipid bilayer or liposome to re-create a membrane like environment for the immunogenic protein.

In accordance with another aspect of the present invention, one or more of the proteins produced from a genetically engineered host (genetically engineered

-8-

with DNA encoding for a 39 kDa T. hyo antigen) may be employed in conjunction with a pharmaceutically acceptable carrier or may be directly conjugated to a carrier or immunostimulant to provide protection against swine dysentery, and in particular swine dysentery induced by T. hyo. The Rotavirus VP6 carrier system developed by VIDO (Veterinary Infectious Disease Organization, Saskatoon, Canada) although not an adjuvant may be a suitable immunostimulant when chemically conjugated to a 39 kDa T. hyo antigen. As hereinabove indicated, such protein(s) is capable of eliciting antibodies which recognize an epitope(s) of one or more of the hereinabove noted 39 kDa T. hyo antigens. Such expressed protein will be sometimes hereinafter referred to as a "recombinant T. hyo antigen," however, as hereinabove indicated, such protein may not correspond to a T. hyo antigen in that it may also be a fragment, derivative or fusion product. The term "recombinant T. hyo antigen" also encompasses such fragments, derivatives and fusion products.

One or more of such 39kDa T. hyo antigens may be employed in the vaccine. In a preferred embodiment, all of the 39kDa T. hyo antigens are employed in formulating a vaccine (i.e., the seven antigens or fragments or derivatives thereof encoded by the seven different T. hyo genes).

The recombinant T. hyo antigen(s) is employed in the vaccine in an amount effective to provide protection against swine dysentery. In general, each dose of the vaccine contains at least 5 micrograms and preferably at least 20 micrograms of such recombinant T. hyo antigen(s). In most cases, the

-9-

vaccine does not include such recombinant T. hyo. antigen in an amount greater than 20 milligrams.

The term "protection" or "protecting" when used with respect to the vaccine for swine dysentery described herein means that the vaccine prevents swine dysentery and/or reduces the severity of swine dysentery.

If multiple doses are given, in general, they would not exceed 3 doses over a six week period.

The vehicle which is employed in conjunction with the recombinant T. hyo. antigen(s) may be any one of a wide variety of vehicles. As representative examples of suitable carriers, there may be mentioned: mineral oil, alum, synthetic polymers, etc. vehicles for vaccines are well known in the art and the selection of a suitable vehicle is deemed to be within the scope of those skilled in the art from the teachings herein. The selection of a suitable vehicle is also dependent upon the manner in which the vaccine is to be administered. The vaccine may be in the form of an injectable dose and may be administered intra-muscularly, intravenously, or by sub-cutaneous administration. It is also possible to administer the vaccine intranasally or orally by mixing the active components with feed or water; providing a tablet form, etc.

Other means for administering the vaccine should be apparent to those skilled in the art from the teachings herein; accordingly, the scope of the invention is not limited to a particular delivery form.

It is also to be understood that the vaccine may include active components or adjuvants in addition to the recombinant T. hyo. antigen or fragments thereof hereinabove described.

-10-

In accordance with a further aspect of the present invention, there is provided an assay for detection or determination of antibody to 39 kDa T. hyo. antigen which employs a 39 kDa T. hyo. protein antigen, of the type hereinabove described, as a specific binder in the assay.

More particularly, there is provided an immunoassay for 39 kDa T. hyo. antibody in which a 39 kDa T. hyo. antigen is employed as a binder, in the assay, for specifically binding 39 kDa T. hyo. antibody.

The assay technique which is employed is preferably a sandwich type of assay wherein the 39 kDa T. hyo. antigen is supported on a solid support, as a binder, to bind 39 kDa T. hyo. specific antibody present in a sample, with the bound antibody then being determined by use of an appropriate tracer.

The tracer is comprised of a ligand labeled with a detectable label. The ligand is one which is immunologically bound by the 39 kDa T. hyo. antibody and such ligand may be labeled by techniques known in the art.

Thus, for example, the 39 kDa T. hyo. antibody bound to the 39 kDa T. hyo. antigen on the solid support may be determined by the use of an antibody for 39 kDa T. hyo. antibody which is labeled with an appropriate detectable label.

In such a sandwich assay technique, the labeled antibody to 39 kDa T. hyo. antibody may be a monoclonal antibody or a polyclonal antibody; e.g. the polyclonal antibody may be anti-swine IgG or may be an antibody which is specific for 39 kDa T. hyo. antibody, which antibody may be produced by procedures known in the art; for example: innoculating an appropriate animal with 39 kDa T. hyo. antibody.

-11-

The detectable label may be any of a wide variety of detectable labels, including enzymes, radioactive labels, chromogens (including both fluorescent and/or absorbing dyes) and the like. The selection of a detectable label is deemed to be within scope of those skilled in the art from teachings herein.

The solid support for the antigen may be any one of a wide variety of solid supports and the selection of a suitable support is deemed to be within the scope of those skilled in the art from the teachings herein. For example, the solid support may be a microtiter plate; a tube, a particle, etc.; however, the scope of the invention is not limited to any representative support. The antigen may be supported on the support by techniques known in the art; e.g., by coating; covalent coupling, etc. The selection of a suitable technique is deemed to be within the scope of those skilled in the art from the teachings herein.

The sandwich assay may be accomplished by various techniques; e.g., "forward"; reverse"; or "simultaneous"; however, the forward technique is preferred.

In a typical procedure, 39 kDa T. hyo. antigen, which is supported on a solid support is initially contacted with a sample containing or suspected of containing 39 kDa T. hyo. antibody to bind specifically any of such antibody present in the sample to such antigen on the support.

After washing of the solid support, the support is contacted with a tracer which binds to 39 kDa T. hyo. antibody. If such antibody were present in the sample, the tracer becomes bound to such antibody bound to such antigen on the solid support, and the

-12-

presence of tracer on the solid support is indicative of the presence of 39 kDa T. hyo. antibody in the sample. The presence of tracer may be determined by determining the presence of the detectable label by procedures known in the art.

Although the preferred procedure is a sandwich assay, it is to be understood that the 39 kDa T. hyo. antigen(s) may be used in other assay techniques, e.g., an agglutination assay wherein the antigen is used on a solid particle such as a latex particle.

In accordance with another aspect of the present invention, there is provided an assay or reagent kit for determining 39 kDa T. hyo. antibody which includes 39 kDa T. hyo. antigen, as hereinabove described, and a tracer comprised of a ligand and a detectable label. The ligand of the tracer is bound by 39 kDa T. hyo. antibody. The reagents may be included in a suitable kit or reagent package, and may further include other components, such as buffers etc. The 39 kDa T. hyo. antigen is preferably supported on a solid support.

DNA fragments may be used as a probe by use of techniques known in the art.

Although the present invention has been particularly described with reference to use of the 39 kDa antigen(s) for imparting protection against T. hyo., one or more of such antigens may be used to produce antibodies (monoclonal and/or polyclonal) by procedures known in the art and such antibodies may be used in a vaccine to impart protection against T. hyo.

Description of Appendices and Drawings

Appendix 1 is a comparison of Gene Products of the 39kDa gene Family without peptide signal sequences from serotype B204 .

-13-

Appendix 2A is the DNA sequence of genes 1-4 encoding antigens 1-4 of the 39kDa gene family from serotype B204.

Appendix 2B is the DNA sequence of genes 5-8 encoding antigens 5-8 of the 39 kDa gene family from serotype B204.

Appendix 3 is the nucleotide sequence of T.hyo gene insert of pTrep 106.

Appendix 4 is a partial DNA sequence of plasmid pTrep 301.

Appendix 5 is predicted amino acid sequences from PCR derived T.hyo. (B204) clones.

Appendix 6 is the predicted protein sequence encoded by pTrep 702.

Appendix 7 is the predicted protein sequence encoded by pTrep 704.

Appendix 8 is the predicted amino acid sequence for pTrep 505.

Figure 1 is a map of the gene family and sub-clones obtained from screening for 39 kDa gene;

Figure 2 is a plasmid map of pTrep 505;

Figure 3 is a schematic of the construction of pTrep 702;

Figure 4 is a schematic of the construction of pTrep 704; and

Figure 5 is a schematic of the construction of the pTrep PCR expression vehicle.

The present invention will be further described with respect to the following examples; however, the scope of the invention is not to be limited thereby. In the Examples, unless otherwise noted, purifications, digestions and ligations are accomplished as described in "Molecular Cloning, a laboratory manual" by Maniatis et al. Cold Spring

-14-

Harbor Laboratory (1982). In the following examples, unless otherwise indicated, transformations are accomplished by the procedure of Cohen et al. PNAS 69 2110 (1973).

EXAMPLE 1 - Purification and Recovery of
Native Antigen

Treponema hyodysenteriae strain B204 was grown in broth culture prepared as follows. Brain/Heart Infusion (Difco) at 37 gms/liter distilled water was autoclaved allowed to cool, then sterile additions were made of a glucose solution (to a final concentration of 5 gm/liter) and fetal calf serum (to final concentration of 5% vol/vol). The media was then prerduced (made anaerobic) by 24 hours of perfusion with a stream of gas composed of 90% nitrogen, 10% carbon dioxide. The complete media was then inoculated with a 1-10% volume of actively growing T. hyo culture, the temperature was maintained at 37°C-39°C, the culture pH was maintained at 6.8, and the culture was continuously perfused with the oxygen free gas (flow rate 50 mls/min/liter of culture).

Cells were removed from the fermentation when they had achieved a density of 5×10^8 /ml or greater (measured by microscopic count). Cells were concentrated by centrifugation then washed and recentrifuged twice in a buffer of 10mM potassium acetate pH 4.75, 150mM potassium chloride. The cells were then resuspended in 10mM potassium acetate pH 4.75 until an optical density of 25-30 (at 600nm) was achieved (as measured on solution dilutions) which is typically about 1/20 the original culture volume.

Extraction method:

-15-

Tween-20 (a non-ionic detergent) was then added to the cell suspension to achieve a final concentration of 0.2%. After gentle agitation for 10 minutes the cells were centrifuged (10,000 xg for 10 min). This supernatant fraction was discarded and the cells were resuspended in an equivalent volume of acetate buffer and then extracted by the addition of Tween-20 to a final concentration of 2.0%. After centrifugation the 2% Tween supernatant (detergent solubilized antigen pool) was saved and the cell pellet was resuspended and re-extracted with Tween-20 in a sequential manner for up to 5 additional cycles with the Tween-20 concentration increasing over the cycles from about 2% up to about 10%. The detergent solubilized supernatant fractions were pooled. This extraction procedure selectively (but not quantitatively) solubilizes surface proteins of T. hyo without lysing or rupturing the bacteria.

To concentrate the antigen preparation, supernatant fractions were subjected to ultracentrifugation (100,000 x g) for 1.5 hours, and the recovered pellet material (HSP) was resuspended in 25mM Tris buffer pH 6.8 and dispersed by sonication.

Antigen Purification

The resuspended HSP was then mixed with 15 volumes of Tris-HCl pH 6.8, 6M urea which had been filtered through a 0.45uM filter. This was stirred at room temperature for several hours. This was centrifuged at 100,000 xg and the supernatant (US1) set aside. The pellet fraction from this step (UP1) was resuspended and extracted with urea a second time. This material was centrifuged as before and the supernatant (US2) and pellet (UP2) were collected.

-16-

The predominant protein constituent of UP2 is 39 KDa protein sometimes referred to as the 39p antigen. The 39p antigen which was further purified by molecular sieve column chromatography in the presence of SDS or electroelution from acrylamide gels.

It is possible to isolate a soluble form of the 39kDa antigen (39s) in addition to the sedimentable form (39p) that is isolated as the major protein component of the urea insoluble pellet (UP2) described above. In order to produce the 39s antigen, T.hyo cells (B204) were extracted with Tween 20 as described above. After the final Tween 20 extraction the residual cell pellet was resuspended in approximately 2ml of 10mM potassium acetate, pH 4.75 per gram wet weight and sonicated. The sonicated cell pellet was separated from the 39s antigen, by centrifugation at 26,000 xg for 15' at 4C. The supernatant was then centrifuged at 100,000 xg for 2 hours at 20°C to pellet any of the sedimentable membrane associated proteins which were also released by sonication. The supernatant (39*), which contains the 39s antigen as its predominant protein component, was then sterile filtered through a .2uM filter and stored at either 4C or frozen. If stored at 4C some proteolytic degradation of the 39s antigen occurs. Additional 39* can be isolated by repeating the above sonication and centrifugation steps on the 26,000 xg cell pellet. Approximately 4mg of 39* can be obtained per liter of original culture volume; a yield roughly equivalent to the yeild of UP2.

The electrophoretic mobility in SDS polyacrylamide gels of the 39s protein and the 39p protein is identical. The two proteins are also

-17-

immunologically cross-reactive. Antisera raised against UP2, or gel purified 39p, will recognize 39s on Western blots. Conversely, antisera raised against 39* will recognize 39p on Western blots. 39S and 39p also comigrate with the predominant protein on the surface of intact T.hyo cells labelled with I^{125} . (Marchalonis, et al. Biochemistry Journal; 24, 921 (1971)). Antisera from swine that have recovered from experimental infections of swine dysentery also recognize either the 39s or 39p form of the 39kDa antigen.

Example 2 - Protein Sequence of 39s and 39 antigens

The fermentation and protein purification are accomplished as in Example 1.

The insoluble material obtained by centrifugation of the second urea extraction (UP2) contains a single major protein component which is 39p antigen. This insoluble protein was solubilized by boiling in 25 mM Tris-HCl pH6.8 containing 3%SDS, 1 mM EDTA, and 70 mM 2-mercaptoethanol. This solution was subjected to gel filtration chromatography over a 30 cm column of Sepharose 6B (from BioRad, Richmond, CA). The 39 kDa peak was identified by gel electrophoresis of column eluant fractions, the appropriate fractions were pooled and the protein concentrated by precipitation with acetone and collected by centrifugation. The pellet was dissolved in 1.1% SDS and then extracted with chloroform/methanol to remove residual SDS.

The amino acid sequence of the amino-terminus of the 39 kDa protein prepared above was determined using sequential Edman degradation in an automated Applied Biosystems gas phase sequenator. The identity of the first 41 amino acids of the protein thus determined are shown below:

-18-

1	10
Met-Tyr-Gly-Asp-Arg-Asp-Ser-Trp-Ile-Asp-Phe-Leu-Thr-His-Gly-	
20	
Asn-Gln-Phe-Arg-Ala-Arg-Met-Asp-Gln-Leu-Gly-Phe-Val-Leu-Gly-	
	41
Asn-Asp-Thr-Ile-Lys-Gly-Thr-Phe-?-?-Arg-	

Amino terminal peptide sequence of 39s was obtained directly from a preparation of 39* which was concentrated by precipitation with acetone and sequenced. Additional internal peptide sequence of the 39s antigen was obtained by digestion with endoproteinase LysC (~1/100 w/w) in 50mM Tris-HCl pH8.5, 0.1% SDS (37°C, 16 hrs.). Proteolytic cleavage products were purified using reverse phase HPLC and sequenced on a Vydac C4 column (250 mm x 4.6 mm, 5µM) developed with a linear gradient of 0%-100B% (0% = 0.1% Trifluoroacetic acid, 100B% = 67% acetonitrile, 33% isopropanol, 0.1% trifluoroacetic acid). Peptide sequence for the 39p antigen extending that found above was determined in a similar fashion. Some additional protein sequence from 39* was also obtained from purified cleavage products after digestion with endoproteinase V8(~1/100 w/w) in 50mM NH₄HCO₃, pH 7.8, 0.1% SDS.

Internal sequences were also determined for the 39p antigen as follows:

An amino acid sequence was determined for an HPLC purified peptide fragment derived from proteolytic digestion of the 39 kDa protein (of the UP2 cell fraction) using endoproteinase Lys-C. 300µg of the 39 kDa protein was first precipitated with acetone and then resuspended in a solution of 4 M urea, 25 mM Tris pH 8.5 and digested with 2.5µg LysC

-19-

(37°C, 16 hrs.). One peptide was obtained as a peak of material eluting off of a C-4 reverse phase column developed with a gradient of acetonitrile, isopropanol (2:1) in 0.1% trifluoroacetic acid.

The purified fragment had the following internal sequence: val gln his ser leu ala trp gly ala tyr ala glu leu tyr val arg pro val gln asp leu glu glu tyr phe glu met asp ile asn. . .

The Amino acid sequence was also determined for the protease resistant component of the 39 kDa component of the UP2 fraction after its digestion with chymotrypsin. A sonicated suspension of UP2 protein at 2mg/ml was incubated at 37 degrees C for 16 hrs. with 20 ug/ml chymotrypsin in a buffer of 25 mM Tris, pH 6.8, containing 0.1% Zwittergent 3-12 detergent. A protease resistant 27 kDa product was isolated by electroelution after preparative gel electrophoresis and precipitated and extracted with chloroform/methanol prior to sequencing. The component had the following sequence:

asp xxx xxx thr lys-

asp tyr met gly ile ser thr asp ile gln leu arg tyr tyr thr xxx
ile asp ala phe asn ala ile arg leu tyr phe lys tyr gly gln xxx
xxx phe

A summary and comparison of the amino acid sequence data obtained from 39p and 39s is found in Table 1 following the Examples. Although the proteins have not been sequenced in their entirety, none of the data identifies any difference in the amino acid sequence of the 39p and 39s antigens. Therefore, one can conclude that the amino acid sequences of the 39p and 39s antigens are very similar and may be identical.

-20-

Example 3 - General Methods

Unless otherwise indicated, the General Methods described below were used in the following examples.

A. Construction of a Genomic Library of T.hyo DNA

48 µg of genomic DNA of T.hyo strain B204 was partially digested with Alu I. EcoRI linkers were kinased with P32 ATP according to manufacturers instructions (Pharmacia LKB Biotechnology, Piscataway, New Jersey) and ligated to the Alu I partially digest T.hyo DNA at a linker concentration of 133 µg/ml using BMB ligase at a concentration of 50 units/ml. Following overnight ligation the ligase was heat inactivated and the reaction was digested with EcoRI.

The DNA was fractionated on an S-200 (Pharmacia) column using 0.3 NaCl, 0.05 M Tris-HCl pH 8.0, 1 mM EDTA, 0.06% sodium azide as a column buffer, in order to remove free linkers and free ATP. The recovered T.hyo DNA was then ligated to dephosphorylated lambda gt11 EcoRI arms obtained from PRomega Biotec (Madison, Wisconsin) and used according to manufacturers specifications. The ligation was then packaged into lambda bacteriophage particles using the in vitro packaging kit, "Gigapack," obtained from Stratagene (San Diego, CA). The phage was then titered on a stationary phage culture of E. coli strain Y1090r-(Promega Biotech). The number of white plaques indicated that the original phage stock contained 1.4×10^7 pfu/ml in a total of 0.5ml.

B. Identification of a Recombinant Phage

In performing mixed oligonucleotide screening for the 39kDa gene, the procedure used was that of W.D. Benton & R.W. Davis Science 196, 180 (1977).

-21-

Duplicate filters were hybridized with each oligonucleotide probe. Approximately 10E6 cpm (1-2 ng probe) of probe was used for filter, overnight at 37C. The hybridization solution consisted of:

- 5X Denharts
- 0.1 μ M rATP
- 250 μ g/ml E. coli tRNA
- 6X NET (1XNET=150mM NaCl, 15 mM Tris-HCl pH 7.5, 1 mM EDTA)
- 0.5% NP40
- 1 mM sodium Pyrophosphate

Prior to hybridization the filters were washed for 2 hours at 37C in hybridization solution. Following hybridization the filters were washed twice at RT (20'/wash) and twice (20'/wash) at 37C in 6X NET, 0.1% SDS and once (20'/wash) at 37C in 6X NET. The filters were then dried and exposed to X-ray film. Positive plaques were selected, rescreened and plaque purified. Phage DNA was isolated using the technique of C. Helms, et al. (DNA 4 39, 1985).

C. PCR Protocol

1. 10ng genomic DNA (B204 or B234) was mixed with 10ul 10x reaction buffer (Perkin Elmer Cetus), 16ul 1.25mM dNTP (each), 25ul primer#1 (4uM), 25 ul primer #2 (4uM) and brought to a final volume of 100ul with Q-H₂O.

2. The mixture was denatured by heating at 94°C for 1.5 min. and annealed at 50°C for 2.5 min. 0.5ul of Taq polymerase (15U/ μ l) (Perkin Elmer Cetus) was added and polymerization was allowed to proceed at 55°C for 10 min.

-22-

3. One more round of denaturation, annealing and polymerization was performed with the time and temperature conditions specified in step #2.

4. Twenty three rounds of denaturation, annealing and polymerization were performed as in step #2 except that the polymerization temperature was increased to 65°C.

5. After the final round of amplification the mixture was extracted with phenol:chloroform (1:1) and chloroform and precipitated with ethanol.

6. After extraction and precipitation the sample was digested with appropriate enzymes (BamH1 and Hind3) and ligated with the desired vector (pUC8 or pUC9).

D. Dot Blot Screening Protocol

1. Grow an overnight culture of bacterial colony to be screened.

2. Spin down 50ul of the overnight culture and remove supernatant.

3. Resuspend the cell pellet in 200ul of 25mM Tris-Cl pH 8.0, 10mM EDTA + hen egg white lysozyme (1mg/ml).

4. Incubate for 5 min. at room temperature.

5. Sonicate briefly (3 seconds).

6. Add 20ul 3N NaOH and incubate 1 hour at 70C.

7. Let cool to room temperature, add 220ul 2M NH₄OAc. Mix.

8. Apply to nitrocellulose filter with aid of vacuum.

9. Let filter air dry. Bake at 90C for 2hrs.

10. Probe filter with desired probe.

E. Preparation of Nick Translated Probe

1. Denature 50ng of the DNA fragment by boiling for 5 min. in 9ul TE. Chill on ice.

-23-

2. Add 5ul gamma dATP (6000 Ci/mMol, 10mCi/ml), 2ul degenerate hexamer in 10x reaction buffer (BMB), 3ul dNTP (25uM dG-,dT-, & dCTP final concentration), 2ul Klenow (2U/ul, BMB).

3. Incubate at 37C for 30 min. Stop with 30ul 10mM EDTA.

4. Separate labeled fragement from unincorporated label with G-50 spin column (BMB).

F. Screening Procedure with Nick Translated Probes

The screening procedure was as follows:

1. Prewash baked nitrocellulose filters in 20mM Tris-Cl pH8.0, 1mM EDTA, 0.1% SDS for 2 hrs. at 37°C.

2. Prehybridize filters for 2 hrs. at 42°C in 50% deionized formamide, 5X Denhardt's, 5X SSPE, 0.1% SDS, salmon sperm DNA (100ug/ml).

3. Denature nick translated probe by boiling for 5 minutes and chilling on ice.

4. Hybridize overnight at 42°C in prehybridization solution and denatured probe.

5. Wash filters 2 times at room temperature in 2X SSC and 0.1% SDS.

6. Wash 1 time in 0.1X SSC at 42°C.

7. Dry filters and expose to x-ray film at -70°C with enhancing screen.

G. Cement Preparation

1. Resuspend cells from an overnight culture in 1/25th original volume in 25mM Tris-Cl pH 8.0, 10mM EDTA + 1mg/ml lysozyme.

2. Incubate 30-60 min. Sonicate to disrupt DNA and reduce viscosity.

-24-

3. Add 1/10th volume 20% Triton X-100. Agitate on lab quake for 2 hours. Sonicate if necessary to reduce viscosity.
4. Centrifuge at 10,000xg for 10 min. to pellet cement.
5. Resuspend cement in 1/25th original volume in 20mM Tris-Cl pH 8.0, 5mM EDTA + 5% Triton X-100. Sonicate. Agitate on lab quake overnight.
6. Centrifuge at 10,000xg for 10 min. to pellet cement. Wash cement with 20mM Tris-Cl pH 8.0, 5mM EDTA. Centrifuge.
7. Resuspend in 1/50th original volume in 20mM Tris-Cl pH8.0, 5mM EDTA.

Example 4 - Identification of the
Gene Encoding the Initial Member of the
39 kDa Antigen family

A set of DNA probes were synthesized using the amino terminal amino acid sequence data shown in Example 2. Each of them were comprised of a pool of degenerate sequences which encompass all the possible combinations of nucleotides which could encode the amino acid sequence of the target region as indicated below. Each probe is 17 nucleotides in length.

-25-

1

Met-Tyr- Gly-Asp-Arg-Asp
 ATG-TAT-GGT-GAT-AGT-GA
 C C C C
 A A
 G G

probe name = COD 555

degeneracy = 128 fold
 (mix of 128 combinations)

10

Trp-Ile-Asp-Phe-Leu-Thr
 TGG-ATT-GAT-TTT-TTT-AC
 C C C C
 A A
 G

probe name = COD 553

degeneracy = 96 fold

18

His- Gly- Asn- Gln- Phe- Arg
 CAT-GGT-AAT-CAA-TTT-AG
 C C C G CC
 A
 G

probe name = COD 556

degeneracy = 128 fold

A lambda GT11 library containing EcoRI linked fragments derived from a partial AluI digest of genomic T.hyo DNA (strain B204 was screened with probes. One phage, 3-5C1 was identified by hybridization to probes 553 and 555. The DNA was examined after digestion with EcoRI and found to contain a 1.6 kb insert.

The Eco RI flanked, 1.6 kb segment of DNA from phage 3-5C1 was isolated by electroelution from an acrylamide gel and then ligated to plasmid pUC 19 which had been linearized by digestion with EcoRI. These DNAs were the ligated together, transformed into E. coli, and a clone containing recombinant plasmid pTrep 106 (Appendix 3) was identified by analysis or restriction digests of plasmid DNA.

Plasmid pTrep 106 was used to direct protein synthesis in an in vitro coupled transcription-translation system containing ³⁵S-Methionine. SDS-gel electrophoresis of the

-26-

protein products of this system showed 39 kDa protein species not seen with the parental plasmid lacking the T. hyo DNA insert. This suggests that the cloned DNA contains the complete coding sequence for the T. hyo 39 kDa antigen and that E. coli is capable of recognizing the treponemal promoter and ribosome binding site and directing the synthesis of this foreign protein.

E. coli strains transformed with plasmid pTrep 106 did not produce significant amounts of the desired 39 kDa T. hyo antigen. Therefore, plasmid construction allowing high level expression of the recombinant antigen was made as follows. The Eco RI flanked, 1.6 kb fragment of pTrep 106 was ligated to plasmid pUC 18 linearized by digestion with Eco RI. The resulting plasmid, pTrep 112, was then cut with PstI and BamHI, then treated with exonuclease III to remove (in a unidirectional manner) the non-coding DNA sequence upstream of the predicted ATG start codon of the 39 kDa T. hyo antigen (Henikoff, Gene 28 p. 351-59 (1984)). At various times during this digestion, DNA aliquots were removed, the exo III inactivated by phenol extraction, the remaining DNA rendered blunt ended by digestion with nuclease S1, and this DNA was then religated and used to transform E. coli. Nucleotide sequencing (Sanger, et al., PNAS 74:5463 (1977)) of plasmid DNA from one such new clone, pTrep 112-1, indicated that a contiguous sequence of 372 codons encoding the mature T. hyo 39 kDa protein and 7 amino acids from the signal sequence were fused downstream of the Hind III site of the parental pUC 18 plasmid. The fusion was in a reading frame to encode a fusion protein whose expression would be regulated by the lac promoter after the orientation of the cloned fragment was

-27-

inverted (see Appendix 4) by cloning into pUC 9 from the HindIII to Eco RI site. E. coli transformed with the resulting plasmid, pTrep 301, produced an insoluble 39 kDa antigen which reacts with sera from swine (both those recovered from swine dysentery as well as animals immunized with the 39 kDa protein purified from T. hyo.) in both an immunoblot and plate ELISA assay.

PURIFICATION OF THE RECOMBINANT FORM OF THE 39 KDA ANTIGEN

E. coli strain CY-15,000 containing plasmid pTrep301 was grown in 250 mls of Luria broth containing ampicillin at 200 µg/ml. The culture was grown for 18 hours at 37° C. The cells were harvested by centrifugation then resuspended in 1/20 their original volume in a buffer of 25 mM Tris, 10 mM EDTA at pH 8.0 and containing lysozyme at 1 mg/ml. After a 30 minute incubation at room temperature the cells had lysed and were then further disrupted by sonication. The non-ionic detergent, Triton X-100 was added to a final concentration of 2%, the cell lysate was mixed for 1 hour at room temperature and then centrifuged at 10,000 xg. The insoluble pellet fraction after these steps was saved. The major protein component of this fraction had a Mr of about 40 kDa as judged by Commassie blue staining of samples after SDS-gel electrophoresis. This same protein component was recognized in Western blot analysis by swine and mouse antisera raised against the authentic 39 kDa T. hyo protein obtained from the UP2 fraction. This recombinant protein was also recognized in immunoblots probed with sera from pigs that had recovered from experimentally induced swine dysentery.

-28-

The predicted amino acid sequence of the 39kDa recombinant antigen obtained in this Example closely resembles but is not identical to the amino acid sequence of the 39 kDa antigen of the UP2 fraction of T. hyo.; however, they have common epitopes recognized in a single sera. As hereinafter indicated, the 39 kDa recombinant produced in this Example corresponds to a protein encoded by gene 1, one of the multiple genes encoding different T. hyo antigens, each having a molecular weight of about 39kDa.

As discussed in Example 4, *infra*, the T. hyo genome contains at least 7 genes encoding related antigens with molecular weights of about 39 kDa. Although the product of only one of these genes is isolated from cells grown in vitro, it is possible that the other members of the gene family are expressed in vivo and are of immunological relevance to protecting against infection in the field. The observation that each of these proteins is preceded by a signal sequence indicates that they will all be exported from the cytoplasm of the cell when expressed. When cells grown in vitro are surface labeled in the presence of I^{125} and lactoperoxidase the 39kDa protein in the KGP fraction is the predominant protein identified. Thus, cells expressing other members of the 39kDa gene family would likely have a much different surface architecture than cells expressing the Copy5 gene in vitro. An immune response mounted against one form of the 39kDa gene family could be only marginally effective against cells expressing one of the other forms.

Example 5
Identification of the genes
encoding additional members of the
39 kDa antigen family

Internal amino acid sequence from 39p was divergent enough from the predicted amino acid sequence of pTrep301 to allow the selection and synthesis of a degenerate oligonucleotide (Cod664, Table 2) that could be used to distinguish between sequences encoding the gene product of pTrep 106 and those encoding the 39 kDa antigen.

The lambda GT11/ B204 library used in the pTrep 106 screening was probed with Cod664 as well as a nick translated probe made from a 411 base pair SphI-BclI fragment encoding the amino terminal portion of the 39kDa protein from pTrep301.

-30-

In addition a library constructed by the ligation of a partial Sau3a digest of B204 genomic DNA and BamHI digested lambda EMBL3 was also screened with these probes. This screening identified a number of phages which hybridized to either Cod664, the nick translated probe 301 Sph-Bcl or both probes.

Hybridizing phage were purified and subcloned into pUC8, 9, 18 or 19 for sequencing and additional manipulations for expression. The recombinant lambda phage, their hybridization patterns, and subclones are elaborated in Table 3 following the Examples.

Based on the DNA sequence from the above subclones and the internal peptide sequences, it was determined that there were at least six related genes encoding similar 39 kDa proteins. Of these six genes, analysis of protein and nucleotide sequence data indicate that gene #5 most likely encodes the 39kDa antigen found in the UP-2 and 39* fractions.

None of the subclones or combinations of subclones contained a full length copy of the #5 gene. Therefore, additional probes were prepared to isolate the remaining portion of the gene encoding the 39kDa antigen of the UP-2 and 39* fractions as well as other genes of the 39kDa family. These probes were based on additional internal sequences of the native antigen as well as DNA sequences of phages of Table 3.

A unique oligonucleotide probe specific for Gene #5 (Cod968) (Table 2) was synthesized and used to screen the GT11 library. Two degenerate oligonucleotides (Cod 1019 and Cod 1020) derived from peptide sequence of a carboxy terminal fragment from the 39s and 39p antigens were also used to screen the GT11 library to obtain a phage(s) which contained more extended coding sequences for the #5 gene. The

-31-

degeneracy of Cod 1019 and 1020 was decreased by assuming that codon usage for some amino acids would be similar to that found in other genes in the 39 kDa family.

Unique oligonucleotides were also synthesized and used to screen the GT11 library for full length genes corresponding to Gene #6 (Cod 934), Gene #7 (Cod 1010 and Cod 1011) and Gene #8 (Cod 1328).

A summary of all of the phages identified through this additional screening, their hybridization patterns and subclones is contained in Table 4.

DNA sequence derived from overlapping subclones indicates that the 39 kDa genes are found in two subfamilies of tandemly repeated 39kDa genes. Family 1 contains (1-4) and Family 2 contains (5-8). Each gene encodes a protein with a presumptive signal sequence directing transport of the protein through the inner bacterial membrane.

The gene sequence and predicted amino acid sequence for each of the full length genes 1-8 is shown in Appendices 2 and 2A. In addition, Figure 1 of the drawings is a map of the gene family and sub-clones obtained from screening the two libraries.

Appendix 1 shows the relationship between the predicted amino acid sequences of processed products of genes 1-7 encoding seven different full length 39 kDa T. hyo antigens as well as the carboxy terminal fragment of gene #8.

The Perkin-Elmer/Cetus polymerase chain reaction system was used as a supplement to screening phage libraries to identify clones containing full length copies of 39 kDa genes. DNA sequence of Genes #1, 2, 3 (only 3') and 4 (only 5') indicated that these genes, although containing unique internal DNA

-32-

sequences, contained identical 5' and 3' DNA sequences. Thus, two linked oligonucleotides corresponding to the 5' sequence (Cod 987), and the reverse complement of the 3' sequence (Cod 988), were synthesized to be used as primers for DNA synthesis. They were then mixed with a template of genomic DNA from either serotype B204 or B234. The oligonucleotide/DNA mixtures were passed through 25 cycles of heat denaturation, annealing and Taq polymerase directed DNA synthesis to amplify genomic DNA sequences between the two oligonucleotide primers. The newly synthesized amplified sequences were digested with Bam HI and Hind III, and cloned into pUC8 or pUC9. If cloned into pUC8 the fragments were oriented in the proper direction and in the proper reading frame for expression from the Lac promoter of a fusion protein comprised of 9 amino acids encoded by the pUC polylinker followed by a full-length copy of the mature forms of the corresponding antigens. Clones were initially screened by the Dot Blot Screening Protocol with unique and discriminating synthetic oligonucleotides derived from clones containing the full-length sequence of Gene #1 (Cod 844), or Gene #2 (Cod 931), or the partial sequence of Gene #3 (Cod 908), and Gene #4 (Cod 932, 1151). In addition the clones were screened with a unique synthetic oligonucleotide which is common to all known forms of the 39kDa gene family (Cod 957). Some clones hybridized only to this nondiscriminant probe and upon DNA sequence analysis were found to correspond to Gene #7 even though there is a slight mismatch of DNA sequence between Cod 987 and the 5' and 3' ends of the gene. A summary of the subclones obtained along with their

-33-

hybridization patterns is found in Table 4 following the Examples.

The PCR system was also used to synthesize and clone the #5 gene encoding the full length 39p/39s antigen. The oligomers used in this procedure were Cod 1054 and Cod 1055. Cod 1054 was derived from the DNA sequence of the 5' end of gene #6, a gene encoding a protein whose amino terminal amino acid sequence is identical to that of the 39s/39p antigen. Cod 1055 was derived from the reverse complement of the DNA sequence encoding the carboxy terminal peptide of the 39s/39p antigen. This sequence distinguishes the #5 gene from any other gene obtained to date. The oligonucleotides were then mixed with genomic DNA from either B204 or B234 and passed through 25 cycles of heat denaturation, annealing and DNA synthesis in the presence of Taq polymerase in order to amplify intervening sequences. The amplified mixture was digested with BamHI and Hind3 and cloned into either pUC8 or 9. Candidate clones were screened for hybridization to Cod 968, Cod 1019, Cod 1020 and Cod 957. One of those clones, pTrep 613, includes the entire coding sequence for gene which is expressed under control of the beta-galactosidase promoter of pUC.

The PCR system was also used to synthesize and clone the #8 gene encoding the full length Copy 8 Antigen. The oligomers used in this procedure were Cod 1359 and Cod 1438, corresponding to the 3' and 5' ends of the gene, respectively. The oligonucleotides were mixed with genomic DNA from either B204 or B234 and passed through 25 cycles of heat denaturation, annealing, and DNA synthesis in the presence of Taq polymerase in order to amplify intervening sequences. The amplified mixture was digested with BamHI and

-34-

SalI and cloned into either pUC8 or 9. Candidate clones were screened for hybridization to Cod 957. One of these clones, pTrep 541, includes the entire coding sequence for the gene which is expressed under control of the beta-galactosidase promoter of pUC.

A schematic of the B204 expression clones derived from the PCR reaction (pTrep 345, 541, 604, 605, 620, 613) is found in Figure 5. The predicted amino acid sequences encoded by these clones are found in Appendix 5.

Expression of recombinant forms of the 39 kDa protein from genomic DNA subclones corresponding to Genes #2 and 6.

pTrep505, a pUC19 based plasmid directs the expression of all but the first 19 amino acids of Gene #2 of the 39 kDa gene family from the Lac promoter. It was constructed from pTrep 323 which contained an EcoRI fragment subcloned from the lambda GT11 library. This EcoRI fragment was subcloned into pWHA142 to place it in the proper orientation and reading frame for expression from the Lac promoter. pWHA142 is a derivative of pUC19 with a GAA reading frame across the EcoRI site. A plasmid map of pTrep505 is presented in Figure 2. The predicted protein sequence from pTrep 505 is presented in Appendix 8.

pTrep 702, a pUC19 based plasmid directs the expression of 13 amino acids from the signal sequence of Gene #6 plus the first 315 amino acids of the mature protein fused to the LacZ complementing peptide in pUC. It was constructed in two steps from pTrep501 and pTrep327 which contain overlapping regions of the #6 gene and share a common BclI site. pTrep501, which contains regions coding for the 5'

-35-

portion of the #6 gene, was digested with BclI and AatII and ligated with BclI-AatII fragment from pTrep 327 which contained the 3' sequences of the #6 gene. The EcoRI fragment from this clone, pTrep701, was then cloned into the EcoRI site of pWHA142 to place the #6 sequence in the proper reading frame for expression from the Lac promoter. A schematic of the construction of pTrep702 is presented in Figure 3. The predicted protein sequence of the recombinant product encoded by pTrep702 is presented in Appendix 6.

An expression clone encoding the full length Copy #6 antigen, pTrep704, was constructed from pTrep702 by replacing its 430 bp NsiI-NdeI fragment with an 847 bp NsiI-NdeI fragment from pTrep508. The 3' cloning site of pTrep508 is downstream of the cloning site contained within pTrep702 and thus contains the DNA sequences encoding the carboxy terminal portion and stop codon of the Copy #6 antigen which are lacking in pTrep702. A schematic of the construction of pTrep704 is presented in Figure 3. The predicted protein sequence of the recombinant product encoded by pTrep704 is presented in Appendix 7.

The recombinant products expressed by pTrep505, pTrep702, and pTrep704 as well as the PCR derived constructs are recovered as insoluble cements in E.coli strain CY15000 after lysis of cells with lysozyme and extraction with Triton X-100 in the presence of EDTA.

These are immunoreactive with sera from animals experimentally infected with T.hyo (B204) and with sera from animals vaccinated with UP2 or the electroeluted 39 kDa protein from UP2.

-36-

The following Table 5 tabulates the 39 kDa expression clones for expressing seven different T. hyo. antigens having a molecular weight of about 39 kDa, with reference to the different genes or copies.

-37-

Table 1 - Amino Acid Comparison of 39s and 39p Antigens

<u>Sequence</u>	<u>Source</u>
MYGDRDSWIDFLTHGNQFRARMDQLGFVLGN?TIKGTFF?R	39p Nterm
MYGDRDSWIDFLTHGNQFRARMDQLGFVL?NGTIKGTFGF??Q?I	39* Nterm
P?S??TK?YMGISTDIQLRYYTGIDAFNAIRLYFKYQGAGFK	39p ^{2,4}
FPYS?STKDYMGISTDIQLRYYTGIDAFNAIRLYFKYQGAGFK	39* ²
TANGASEYFAQSLGFEARFYFLNTPVGNVTINPFIKVVNTA	39p ²
TANGASEYFAQSLGFEARFYFLNTPVGNVTINPFIKVVNTAL	39* ²
AQAVLGITANSDVVSLYVEPSLGYQATYLGK	39p ²
AQAVLGITANSDVVSLYVEPSL	39*
HISENPYLNIDSK	39p ¹
HISENPYLNIDSK	39* ²
VQHSLAWGAYAELYVRPVQDLE?YFEMDIN	39p ¹
RNGVPVNFATSTGIT?YLPALGG?Q	39p ²
MDINNSDSKRNGVPVNFATSTGITWYLPALGGAQ	39* ³

Notes--All sequence derived from fragments generated by digestion with LysC in the presence of SDS unless otherwise indicated.

¹LysC digest in absence of SDS

²Fraction gave multiple sequences which were resolved on basis of intensity and DNA predicted amino acid sequence

³GluC digest in presence of SDS

⁴Sequence begins with residue #2 due to machine failure and loss of residue #1

Single amino acid code used above is as follows:

A=ala	H=his	P=pro	W=trp
C=cys	I=ile	Q=gln	Y=tyr
D=asp	K=lys	R=arg	
E=glu	L=leu	S=ser	
F=phe	M=met	T=thr	
G=gly	N=asn	V=val	

Table 2

COD	Sequence	Source of Sequence	Specifi.
664	ACG-AAG-GAT-TAT-ATG-GG A A C C T C	39p internal peptide sequence	for Copy 5, 6
844	TTAATCCGCATGATA	pTrep 106	
908	GTTTCATCACAAGCAAA	pTrep 333	3, 4
931	ATGAATATGACGGATAA	pTrep 330	2
932	AAAGTTGATAAACAAGG	pTrep 333	4
934	TATCATCCTTCTAATCCT	pTrep 331	6
957	CCGAAAGTACCTTTAAT	pTrep 106	ALI
968	TATAATCCTTATGATCCT	pTrep 317	
987	GAATTCCGGATCCATGTATGG- AGATCAGGACGATTGGATT	pTrep 106	1-4
988	GTCGACAAGCTTATAATTAAA- ATTCTGGCAAATACCAAGT	pTrep 106	1-4
1010	ATATTGACTGATAGTAT	pTrep 506	
1111	AAATAATTTTGATATG	pTrep 506	
1020	GAT-AGA-AAA-AGG-AAT-GG TCT A C C	39* internal peptide sequence	6
1019	AAA-AGG-AAT-GGA-GTG-CC A C T A T C	39* internal peptide sequence	6
1054	GAATTCCGGATCCATGTATGGCGACAG- AGATTCTTGGATC	pTrep 326	5, 6
1055	GTCGACAAGCTTATAATTATTGAGCAC- CGCCTAAAGCAGG	pTrep 510	5
1092	AGTATGTTTGAACCAATA	pTrep 333	3
1151	TCATATGTATCGTGTATA	pTrep 333	4
1095	GGAGTACCTAAACTTCAA	pTrep 337	8
1248	CGACAGAGATTCTTGA	pTrep 613	5, 6

Table 2 (continued)

COD	Sequence	Source of Sequence	Specific for Copy #
1328	GAATTCAATTACGGATT	pTrep 537	n
1359	GTCGACCTGCAGTTATTA- TTGTAAAGCAGGTAAATA- CCA	pTrep 520	R
1438	GAATTCGGATCCATGTATGG- TGCAGACAACACATGGCTT	pTrep 537	R

-40-

Table #3

Table #3--Identification of recombinant phage
and subclones

Identified by Screening with			
Phage*	pTrep	301Sph-Bcl	Cod664
9B^	317		+
54-3	323	+	
53-3B	326	+	+
52-1A	327	+	+
55-1	329	+	
56-3	330	+	
51-3	333	+	
51-1B	331	+	+
53-1B	337	+	+

* All recombinant phage are from the GT11 library unless otherwise noted

^ From EMBL3 library

Table #4

Table #4--Identification of recombinant phage
and subclones

Identified by Screening with

Phage	pTrop	301Sph-Bcl	908	934	968	1010	1019	1020	1092	1095	1151	1248	1328	157
21	506	+		+										
17	509	+		+										
130	517	+												
544	525		+			+			+					
163	510							+						
2	631						+							
519	520									+				
109	515				+									
441	349											+		
101A	537												+	+

Table 5
Expression Clones from PCR

Clone	Serotype	Cod957	Cod844	Cod908	Cod931	Cod932	Cod968	Cod1151
pTrep345	B204	+						+
pTrep605	B204	+			+			
pTrep604	B204	+		+		+		
pTrep608	B234	+			+			
pTrep609	B234	+				+		
pTrep610	B234	+	+					
pTrep613	B204	+					+	
pTrep620	B204	+						
pTrep651	B234	+						
pTrep541	B204	+						

-43-

Table 6

<u>Clone</u>	<u>Copy#</u>	<u>Source</u>	<u>Comments</u>
pTrep301	1	genomic	
pTrep505	2	genomic	lacks amino terminus
pTrep604	3	pcr	
pTrep345	4	pcr	
pTrep613	5	pcr	
pTrep702	6	genomic	lacks carboxy terminus
pTrep704	6	genomic	full length
pTrep620	7	pcr	1 amino acid substitution at C terminus relative to genomic sequence
pTrep605	2	pcr	full length
pTrep541	8	pcr	full length

Example 6

Use of 39* Antigen in a Vaccine

In two vaccination studies, purified 39* which includes 39s antigen and which is prepared according to the procedure of Example 1 (note pages 11 and 12) was tested in comparison with no challenge, vaccination with adjuvant, or vaccination with the commercially available Hyguard (Haver Labs) product.

In the first study, six pigs per test group were used. The pigs averaged 22.6 lb and were approximately 5-6 weeks of age. Five groups of pigs were given two doses, the first on day 0 and a booster on day 36. The injections were given I.M., in the neck with 1 mg/dose of native antigen. Animals were challenged on day 50 by stomach intubation using a pure culture of T. hyo. (B204) at 5.5×10^2 cfu per pig. The study was terminated on day 92.

Vaccines were given with Emulsigen adjuvant. Emulsigen was used as an adjuvant control, mixed with Dulbecco's PBS buffer. The Hyguard, a bacterin, was administered according to manufacturer's directions.

Animals were monitored daily for clinical signs of swine dysentery. Microbiological evaluation of routine weekly rectal swabs was conducted for T. hyo. and Salmonella. Animals showing signs of bloody diarrhea were swabbed and evaluated on that day. Weekly postchallenge pigs were weighed and their feed intake determined. The experimental results are shown below.

-46-

Example 7Use of Native Antigen

A second study was conducted to confirm the results of Example 5 where the 39* vaccinates performed exceptionally well when challenged intra-gastrically with a measured dose of T. hyodysenteriae. As shown below, 1 of 9 animals vaccinated with a 1 mg dose of 39* and challenged by stomach intubation with 4.4×10^2 CFU of T. hyo (B204) developed clinical signs of swine dysentery in comparison to 3 of 6 adjuvant vaccinates and 0 of 6 HyGuard vaccinates.

<u>Vaccinate Group</u>	<u># of Breaks/ # Animals</u>	<u>Day of Onset</u>
No Challenge	0/6	
Adjuvant	3/6	11, 11, 12
HyGuard	0/5	
39*, 1mg	1/6	29
39*, 1mg--new lot	0/3	

Example 8Production of Antibody

A cement preparation (Part G of Example 4) in an amount of 25-200 µg in Emulsigen is intermuscularly injected into pigs. Two weeks later, the pig is boosted with an identical dose. Two weeks after the boost, the pig is bled and the blood is allowed to clot. Immune serum is separated by centrifugation at 4°C.

Numerous modifications and variations of the present invention are possible in light of the above teachings; therefore, within the scope of the

-47-

appended claims, the invention may be practiced otherwise than as particularly described.

-45-

Appendix 1

COPY#

AA #

6 ---R-S---H---S---S-AIG-A-N-ISE-TENV
 7 ---R-S---I---S---S-AIG-A-N-ISE-TENV
 8 ---L---N---DAV-SL-S-ISEKTDL
 1 MYGDDDDWIDFLTDGNGFRARMDELGFVLGNSTIKSTFGFATGSESTQLGYILLNNLST
 2 ---S---L---ATYKCY-
 3 ---F---S---ATYKCY-
 4 ---F---N---S-L---

60

5 E-QT---A---F-G-V--- YVNPRL-V---N---AV---
 6 E-QT---A---F-G-V--- YVNPRL-V---N---AV---
 7 G-D---V-M---DI-S-L--- YNSTL-V---V---N---A---I-
 1 YLGATISGGIGYTSEAFSIGI8YNTSHSLFPTSDNFGSHTPVLMINALNDNLRIIVFVG
 2 ---L---TPL-I-Y---
 3 ---TPL-I---
 4 ---F-T---C---TPL-I---

120

8 ---S-HDPFNDSAKFPYSSSTKD-M---I-L---A---F---A-F---A-SAS
 9 ---S-HDPFNDSAKFPYSSSTKD-M---I-L---A---F---A-F---A-G
 7 IAAEKDPFGK YTISQYK---E---V---QV---I---S---NKKMF
 1 ILVHNESIDG LSYRDNYLGISTGTQIRYTGIDAFNEIRLVVYBQLSYKINPHDTINY
 2 ---DGNMMH TONIN-LY-F---SSSYTDS-
 3 ---Y-GNVGKVDKQGNIS-SHD---NA-YVGK--
 4 ---Y-GN-QKVDKQGNIHDTYD---I---NA-YVGK--

180

8 -T-YQ-QTIS-QL-LH---VL-N---V---L-VDFAS TV-AKGKGNVFPATAFD-RL
 6 AS-YF-Q-L---A---P---VY---K-V-RT---GEAVONTVS--H
 5 AS-YF-Q-L---A---P---VY---K-V-RT---GEAVONTVS--N
 7 DMF-Q---L---RTI---S---ASSDV---GESLVNT
 1 TQEVLAASFGBFETRFYFLNTAVG NVTINPFIKWAYNTALHGYSTMVRALDSMYEEIESYY
 2 DE-FF---F---S---SV!--
 3 EE-FFS---F---I---S-F-P---S
 4 EE-LFS---T---V6---I---T-LQP--J--

240

6 TAWX- NAWAD-PHSIY-REL--LKIIPSVLSV-T-Y-N-IF-G--- VQD--
 5 -SN-NYKLDAGFR-IGKDF-S---KA---I---V---Q-T---NIS-N-Y L
 5 -YD-NYKLDAGFR-IGKDF-S---KA---I---V---Q-T---NIS-N-Y L
 7 T-SKKEV-E---N-AA---L---ML---N-V-K---YKDSK
 1 FDSPA QSYEDINVKWDKNPYDVTQAVLGVANSDDVSLYVEPSLGYRAKYLSKLTIEDPDGKV
 2 S-RT- KAABN-A---
 3 S-R-V S-QAN-A---
 4 -R-V S-EV-JY-L---Y---M Q-E--

300

8 VKGS-LT-T-Y-Q-B-I-R-B---GVPKLBQNPISGNSM--V-GAN---ALQ
 5 -I-S-Q-S-A---VR-Q---GGTROES---Y-K---A-N 365
 5 -I-S-Q-S-A---VR-Q---I---KRN -V-N-TS---ALGGAC 367
 7 Y-G-N-Y---Q-I---I---GN-RQ-SS--IY-E---L 369
 1 NFDKVNHYLSWGAYAELYITPVKOLENYFEMDVNN SCSGST GIPVSFASTT3ITWYLPF 365
 2 ---C--- 365
 3 -L--- 367
 4 -L--- 365

-49-

Appendix 2 A

FIGURE 8 --DNA Sequence of B204 Genes Encoding 39 kDa Antigens 1-4

1 GAATTCATCAGAAATATTTTTTATAATATAAATATTTTTTTATCTTCTATATCTATA
 60 TATTTGCTGAGATATCTTTTAATACATAATAANGCATTTTTTTNCATTNATTTTCANTT
 120 TGTAAATCATATATAATATTGTTAATATCATTAAATTAAATTATCAATATTTAAATTATT
 180 ACTATCTTGTTTTTTTATTTTTTATCATAGAT9AAGCACCAGCAGCTCCTATAGCCAATG
 240 GTATACCTATCACAGGCAATAAGGTAGCTCCTAATACAGCACTAGTTGCTGCTGCCGTTG
 300 CACCTATAGCACCAGCTGCTACTGCAGTACCGCCTAATAAACTAGAAATATAAAGATGGTA
 360 TTTGTTTTATTGATTCATTATTTTAGATAAATCTTGCAATCAAAATATGTTAAGTTTTTTA
 420 AAATAGGTATGAATTTTGATGAATATATTTCTTCTCTATATTTTTTGCTATTAAAC
 480 AAGAAAATAAAATTTTCTATTGTGCTTTTCCAGTGTTATTCTCTCCGGCTATTACAG
 540 TAATACCATCTATTGTTATATTAGCTTCTTTGATTTTAGAAAAATTTTTATATTCAATT
 600 CCATTATATTTTATCCTTTAATAATTCTGTTATTATTATAACATAAAAAATATTAATAA
 660 AAATATATTAATTTTATTAAATCTAATTCTTGAGCTATTTTATATTTTAGTATAATAA
 * Presumed translational start
 MetLysLysPhePheLeuIleMetThrVal
 720 AAAATATAAACTTCAATTTAGGTATGTATAATGAAAAAGTTTTTCTAATTATGACAGTA
 of Copy 1
 LeuLeuSerMetSerTyrCysSerIlePheGlyMetTyrGlyAspGlnAspAspTrpIle
 780 TTATTAAGTATGTCATATTGTTCAATTTTGGTATGTATGGAGATCAGGACGATTGGATT
 AspPheLeuThrAspGlyAsnGlnPheArgAlaArgMetAspGlnLeuGlyPheValLeu
 840 GATTTTCTTACAGACGGCAATCAGTTTAGAGCTAGAAATGGATCAATTAGGATTTGTTTTA
 GlyAsnSerThrIleLysGlyThrPheGlyPheArgThrGlnSerSerSerThrGlnLeu
 900 GGTAAATAGTACTATTAAAGGTACTTTTCGGTTTTAGAACTCAAAGTTCATCAACTCAATTA
 GlyTyrIleLeuLeuAsnAsnAsnLeuGlyThrTyrLeuGlyAlaThrIleSerGlyGly
 960 GGATATATTTTGTGAATAATAATCTTGGTACTTATTGGGAGCAACTATTTCTGGCGGT
 IleGlyTyrThrSerGluAlaPheSerIleGlyIleGlyTyrAsnTyrThrSerHisSer
 1020 ATAGGATATACTTCTGAGGCTTTTAGTATAGGCATAGGCTATAATTATACCAGCCATTCC
 LeuPheProThrSerAspAsnPheGlySerHisThrProValLeuMetIleAsnAlaLeu
 1080 TTATTTCTTACTAGCGATAACTTTGGTTCTCATACTCCAGTACTTATGATTAATGCTTTA
 AsnAspAsnLeuArgIleValIleProValGlnIleLeuValHisAsnGluSerIleAsp
 1140 AATGATAATTTGAGGATAGTTATTCCTGTGCAATATTAGTACATAATGAAAGTATTGAT
 GlnLeuGlyTyrTyrArgAspAsnTyrLeuGlyIleSerThrAspThrGlnIleArgTyr
 1200 CAACTTGGTTACTATAGAGATAATTATTAGGTATAAGTACTGATACGCAATAAGATAT
 TyrThrGlyIleAspAlaPheAsnGluIleArgLeuTyrValLysTyrGlyGlnLeuGly
 1260 TATACAGGCATAGATGCTTTTAATGAAATAAGATTATATGTAATAATATGGCAATTAGGA

-50-

TyrLysIleAsnProHisAspThrIleAsnTyrThrGlnGluValLeuAlaArgSerPhe
1320 TATAAAATTAATCCGCATGATACTATAAAATTATACACAAGAAGTTTTAGCAAGATCATT

GlyPheGluThrArgPheTyrPheLeuAsnThrAlaValGlyAsnValThrIleAsnPro
1380 GGTTTTGAAACAAGATTCTATTTTTGAATACTGCTGTTGGAAATGTAAGTATCAATCCT

PheIleLysValAlaTyrAsnThrAlaLeuHisGlyTyrSerThrMetValArgAlaLeu
1440 TTTATTAAAGTAGCATATAATACAGCTTTGCATGGATATAGTACCATGGTAAGAGCATTG

AspGlyMetTyrGluGluIleGluGlyTyrTyrProAspSerProAlaGlnSerTyrGlu
1500 GATGGTATGTATGAAGAAATAGAAGGTTATTATCCAGATAGTCCTGCTCAATCATATGAA

AspIleAsnValLysTrpAspLysAsnProTyrAspValThrValGlnAlaValLeuGly
1560 GATATTAATGTTAAATGGGATAAGAATCCTTATGATGTAAGTGTGCAGGCAGTATTGGGA

ValThrAlaAsnSerAspIleValSerLeuTyrValGluProSerLeuGlyTyrArgAla
1620 GTAACTGCTAATAGCGATATAGTATCACTTTATGTTGAGCCTTCTTTAGGTTATAGGGCT

LysTyrLeuGlyLysLeuThrTyrGluAspProAspGlyLysValAsnPheAspPheLys
1680 AAATATTTAGGAAAATTAACATATGAAGATCCAGATGGAAAAGTTAATTTTGATTTTAA

ValAsnHisTyrLeuSerTrpGlyAlaTyrAlaGluLeuTyrIleThrProValLysAsp
1740 GTTAATCATTATTATCTTGGGGTGCTTATGCAGAGCTTTATATAACACCGGTAAGAT

LeuGluTrpTyrPheGluMetAspValAsnAsnSerAspSerAspSerThrGlyIlePro
1800 TTAGAATGGTATTTTGAAATGGATGTTAATAATAGTGATTACAGATTCTACAGGTATACCT

ValSerPheAlaSerThrThrGlyIleThrTrpTyrLeuProGluPheOC
1860 GTTAGTTTTGCTTCTACTACAGGAATAACTTGGTATTTGCCAGAAATTTTAATTATAAAGC

1920 AAATTTTATATGATAAAATAAAAAATGTGGGGTATTTATTATTAAAAATAAATACCCCA

1980 CATTTTATTAAATAATTTTTTCAGTAATTTTACATTATATATTTTTTAGTATAATAAAA

* Presumed translational start of
MetLysLysIlePheLeuIleMetThrValLeu
2040 ATATAAACTTAATTTTAGGTATATACAATGAAAAAATTTTTCTAATTATGACAGTATTA

Copy 2
LeuSerMetSerTyrCysSerIlePheGlyMetTyrGlyAspGlnAspAspTrpIleAsp
2100 TTAAGTATGTCATATTGTTCAATATTTGGTATGTATGGAGATCAGGACGATTGGATTGAT

PheLeuThrAspGlyAsnGlnPheArgAlaArgMetAspGlnLeuGlyPheValLeuGly
2160 TTTCTTACAGACGCCAATCAGTTTAGAGCTAGAATGGATCAATTAGGATTTGTTTAGGT

AsnSerThrIleLysGlyThrPheGlyPheArgSerGlnSerLeuSerThrGlnLeuGly
2220 AATAGCACCAATTAAAGGTACTTTCCGTTTTAGATCTCAGAGTTTATCAACTCAATTAGGA

TyrIleLeuAlaIleTyrLysAspTyrThrTyrLeuGlyAlaThrIleSerGlyGlyIle
2280 TATATTTTGGCTATATATAAAGATTATACTTATTTAGGAGCAACTATTTCCGGCGGTATA

GlyTyrThrSerGluAlaPheSerIleGlyLeuGlyTyrAsnTyrThrThrProLeuPro
2340 GGATATACTTCTGAGGCTTTTAGTATAGGTTTAGGTTATAATTATACTACACCGCTTCTCT

IleSerTyrAsnPheGlySerHisThrProValLeuMetIleAsnAlaLeuAsnAspAsn
2400 ATTAGTTATAACTTTGGTTCTCACTCTGTTACTTATGATTAATGCTTTAAATGATAAT

-51-

LeuArgIleValIleProValGlnIleLeuValHisAspGlyAsnMetAsnMetThrAsp
2460 TTGAGGATAGTTATTCCTGTACAAATATTAGTACATGATGGTAATATGAATATGACGGAT

AsnIleAsnTyrLeuTyrAsnPheLeuGlyIleSerThrAspThrGlnIleArgTyrTyr
2520 AATATTAAATTATTTATATAATTTTTTAGGTATAAGTACTGATACTCAAATAAGATATTAT

ThrGlyIleAspAlaPheAsnGluIleArgLeuTyrValLysTyrGlyGlnLeuGlyTyr
2580 ACAGGCATAGACGCTTTTAATGAAATAAGATTATATGTAAATACGGACAATTAGGATAT

LysGlyGlySerTyrThrAspLysSerTyrAspGluGluPhePheAlaArgSerPheGly
2640 AAAGCGGTTTCATATACGGATAAAAGTTATGATGAAGAATTTTTTGCAGATCATTGGT

PheGluThrArgPheTyrPheLeuAsnThrAlaValGlyAsnValThrIleAsnProPhe
2700 TTTGAACAAGATTCTATTTTTGAATACTGCTGTTGGAAATGTAACATCAATCCTTTT

IleLysValAlaTyrAsnThrAlaLeuHisGlyPheSerThrMetValArgSerLeuAsp
2760 ATTAAAGTAGCATATAATACAGCTTTCATGGATTAGTACTATGGTAAGATCATTAGAT

SerValIleGluGluIleGluGlyTyrSerSerAspArgThrAlaLysAlaAlaGlyAsn
2820 AGTGTCAATTGAAGAAATAGAAGGTTATAGTTCAGATCGTACCGCTAAGCAGCAGGAAAT

IleAsnAlaLysTrpAspLysAsnProTyrAspValThrValGlnAlaValLeuGlyVal
2880 ATTAATGCTAAATGGGATAAGAATCCTTATGATGTAACGTGTCAGGCAGTATTGGGAGTA

ThrAlaAsnSerAspIleValSerLeuTyrValGluProSerLeuGlyTyrArgAlaLys
2940 ACTGCTAATAGCGATATAGTATCACTTTATGTTGAGCCTTCTTTAGGTTATAGGGCTAAA

TyrLeuGlyLysLeuThrTyrGluAspProAspGlyLysValAsnPheAspPheLysVal
3000 TATTTAGGAAAATTAAACATATGAAGATCCAGATGGAAGTTAATTTTGATTTTAAAGTT

AsnHisTyrLeuSerTrpCysAlaTyrAlaGluLeuTyrIleThrProValLysAspLeu
3060 AATCATTATTTATCTTGGTGTGCTTATGCAGAGCTTTATATAACACCTGTAAAGATTTA

GluTrpTyrPheGluMetAspValAsnAsnSerAspSerAspSerThrGlyIleProVal
3120 GAATGGTATTTTGAAATGGATGTTAATAATAGTGATTCTACAGGTATACCTGTT

SerPheAlaSerThrThrGlyIleThrTrpTyrLeuProGluPheOC
3180 AGTTTTGCTTCTACTACAGGAATAACTTGGTATTTGCCAGAAATTTTAATTATAAAGCAA

3240 TTTTATATGACAAAATAAAAAATGGGCATTTATTATTAAAAAATAAATACCCACATT

3300 TATTAAATAACTTCTTAAATAATTTTACA4TTTATATTTTATTAGTATAATAAAATATAA

* Presumed translational start of Copy 3
MetLysLysSerPheLeuIleMetThrValLeuLeuSer

3360 AGTTAAATTTAGGTGTGTACAATGAAAAAAGTTTTCTAATTATGACAGTATTATTAAGT

MetSerTyrCysSerIlePheGlyMetTyrGlyAspGlnAspAspTrpIleAspPheLeu
3420 ATGTCATATTGTTCAATATTTGGTATGTATGGAGATCAGGACGATTGGATTGATTTCTT

ThrAspGlyAsnGlnPheArgAlaArgMetAspGlnPheGlyPheValLeuGlyAsnSer
3480 ACAGACGGTAATCAGTTTAGAGCTAGAATGGATCAATTTGGATTCTTTTAGGTAAATAGC

ThrIleLysGlyThrPheGlyPheArgSerGlnSerLeuSerThrGlnLeuGlyTyrIle
3540 ACCATTAAAGGTACTTTTCGGTTTTAGATCTCAGAGTTTATCAACTCAATTAGGATATATT

LeuAlaIleTyrLysAspTyrThrTyrLeuGlyAlaThrIleSerGlyGlyIleGlyTyr
3600 TTGGCTATATATAAAGATTATACTTATTTAGGAGCAACTATTTCCGGCGGTATAGGATAT

-52-

ThrSerGluAlaPheSerIleGlyLeuGlyTyrAsnTyrThrThrProLeuProIleSer
3660 ACTTCTGAGGCTTTTAGTATAGGTTAGGTTATAATTATACTACACCGCTTCCTATTAGT
AspAsnPheGlySerHisThrProValLeuMetIleAsnAlaLeuAsnAspAsnLeuArg
3720 GATAACTTTGGTTCTCATACTCCTGTACTTATGATTAAATGCTTTAAATGATAATTGAGG
IleValIleProValGlnIleLeuValTyrAsnGlyAsnValGlnLysValAspLysGln
3780 ATAGTTATTCTGTACAAATATTAGTATATAATGGTAATGTTCAAAAAGTTGATAAACAA
GlyAsnIleSerTyrSerHisAspTyrLeuGlyIleSerThrAspThrGlnIleArgTyr
3840 GGTAAATATCTCTTATTACATGATTATTAGGTATAAGTACTGATACGCAATAAGATAT
TyrThrGlyIleAspAlaPheAsnGluIleArgLeuTyrValLysTyrGlyGlnLeuGly
3900 TATACAGGTATAGATGCTTTTAAATGAAATAAGATTATATGTAATAATATGGCAATTAGGA
TyrLysAsnAlaProTyrValGlyLysAsnTyrGluGluGluPhePheSerArgSerPhe
3960 TATAAAAATGCTCCGTATGTTGGTAAGAATTATGAAGAAGAATTTTTTCAAGGTCATTT
GlyPheGluThrArgPheTyrPheLeuAsnThrAlaValGlyAsnValThrIleAsnPro
4020 GGTTTTGAACAAGATTCTATTTTTTGAATACTGCTGTTGGAAATGTAACATATCAATCCT
PheIleLysValAlaTyrAsnThrAlaLeuHisGlyPheSerThrMetIleArgAlaLeu
4080 TTTATTAAAGTAGCATATAATACAGCTTTGCATGGATTTAGTACCATGATAAGAGCATT
AspSerMetPheGluProIleGluGlyTyrSerSerAspArgProValSerSerGlnAla
4140 GATAGTATGTTTGAACCAATAGAAGGTTATAGTTCAGATCGTCTGTTTCATCACAAAGCA
AsnIleAsnAlaLysTrpAspLysAsnProTyrAspValThrValGlnAlaValLeuGly
4200 AATATTAAATGCTAAATGGGATAAGAATCCTTATGATGTAAACAGTGCAGGCAGTATTGGGA
ValThrAlaAsnSerAspIleValSerLeuTyrValGluProSerLeuGlyTyrArgAla
4260 GTAACCGCTAATAGCGATATAGTATCACTTTATGTTGAGCCTTCTTTAGGTTATAGGGCT
LysTyrLeuGlyLysLeuThrTyrGluAspProAspGlyLysValAsnLeuAspPheLys
4320 AAATATTTAGGAAAATTAACATATGAAGATCCAGATGGAAAAGTTAATTTGGATTATAA
ValAsnHisTyrLeuSerTrpGlyAlaTyrAlaGluLeuTyrIleThrProValLysAsp
4380 GTTAATCATTATTATCTTGGGGGGCTTATGCAGAGCTTTATATAACACCGTAAAAGAT
LeuGluTrpTyrPheGluMetAspValAsnAsnSerAspSerAspSerThrGlyIlePro
4440 TTGGAATGGTATTTTGAAATGGATGTTAATAATAGTGATTTCAGATTCTACAGGAATACCT
ValSerPheAlaSerThrThrGlyIleThrTrpTyrLeuProGluPheOC
4500 GTTAGTTTTGCTTCTACTACAGGAATAACTTGGTATTTGCCAGAATTTTAATTATAAAGC
4560 AAATTTTATATGACAAAATAAAAAATGNGGGTATTTATTATTAAAAAATAAATACCCCG
4620 TTTTTTATTAAATAACTTCTTAAATAATTTTACATTTTTTATTATTATTAGTATAATAAA

* Presumed translational start of
MetLysLysPheLeuIleMetThrValLeuLeu

4680 ATATAAAGTTAAATTTAGGTGTGTACAATGAAAAAGTTTCTAATTATGACAGTATTATTA

Copy 4

SerMetSerTyrCysSerIlePheGlyMetTyrGlyAspGlnAspAspTrpIleAspPhe
4740 AGTATGTCATATTGTTCAATATTTGGTATGTATGGAGATCAGGACGATTGGATTGATTTT

-53-

LeuThrAspGlyAsnGlnPheArgAlaArgMetAspGlnPheGlyPheValLeuGlyAsn
4800 CTTACAGACGGTAATCAGTTTAGAGCTAGAATGGATCAATTTGGATTTCGTTTTAGGTAAT

AsnThrIleLysGlyThrPheGlyPheArgSerGlnSerLeuSerThrHisLeuGlyTyr
4860 AACACCATTAAAGGTACTTTTCGGTTTTAGATCTCAGAGTTTATCAACTCACTTAGGCTAT

IleLeuLeuAsnAsnAsnPheGlyThrTyrPheGlyThrThrIleSerCysGlyIleGly
4920 ATTTTGTAAATAATAATTTTGGTACTTATTTTGAACAACATATATCATCGGTATAGGA

TyrThrSerGluAlaPheSerIleGlyIleGlyTyrAsnTyrThrThrProLeuProIle
4980 TATACTTCTGAGGCTTTTAGTATAGGTATAGGTATAATTATACTACACCGCTTCCTATT

SerAspAsnPheGlySerHisThrProValLeuMetIleAsnAlaLeuAsnAspAsnLeu
5040 AGTGATAACTTTGGTTCTCATACTCCTGTACTTATGATTAATGCTTTAAATGATAATTTG

ArgIleValIleProValGlnIleLeuValTyrAsnGlyAsnIleGlnLysValAspLys
5100 AGGATAGTTATTCCTGTACAAATATTAGTATATAATGGTAATATTCAAAAAGTTGATAAA

GlnGlyAsnIleHisAspThrTyrAspTyrLeuGlyIleSerThrAspThrGlnIleArg
5160 CAAGGTAATATACACGATACATATGATTATTTAGGTATAAGTACTGATACGCAATAAGA

TyrTyrThrGlyIleAspAlaPheAsnGluIleArgLeuTyrIleLysTyrGlyGlnLeu
5220 TATTATACAGGTATAGATGCTTTTAATGAAATAAGATTATATATAAAATATGGACAATTA

GlyTyrLysAsnAlaProTyrValGlyLysAsnTyrGluGluGluLeuPheSerArgSer
5280 GGATATAAAAATGCTCCGTATGTTGGTAAAAATTATGAAGAAGAACTTTTTTCAAGGTCA

PheGlyPheGluThrArgPheTyrPheLeuAsnThrThrValGlyAsnValThrIleAsn
5340 TTTGGTTTTGAAACAAGATTCTATTTTTTGAATACTACTGTTGGAAATGTAATATTAAT

ProPheIleLysValAlaTyrAsnThrAlaLeuHisGlyValGlyThrMetIleArgAla
5400 CCTTTTATTAAAGTAGCATATAATACAGCTTTGCATGGAGTTGGTACTATGATAAGAGCA

LeuAspThrMetLeuGlnProIleGluAspTyrTyrProAspArgProValSerSerGln
5460 TTAGATACTATGCTTCAACCAATAGAAGATTATTATCCAGATCGTCCTGTTTCATCACAA

ValAspIleAspTyrLysLeuAspLysAsnProTyrAspValThrValGlnAlaValLeu
5520 GTAGATATTGATTATAAATTGGATAAGAATCCTTATGATGTAAGTGTGCAGGCAGTATTG

GlyValThrAlaAsnSerAspIleValSerLeuTyrValGluProSerLeuGlyTyrLys
5580 GGAGTAACCGCTAATAGTGATATAGTATCACTTTATGTTGAGCCTTCTTTAGGTTATAAG

AlaLysTyrLeuGlyLysMetGlnAspGluLysValAsnLeuAspPheLysValAsnHis
5640 GCTAAATATTTAGGAAAAATGCAAGATGAAAAAGTTAATTTGGATTTTAAAGTTAATCAT

TyrLeuSerTrpGlyAlaTyrAlaGluLeuTyrIleThrProValLysAspLeuGluTrp
5700 TATTTATCTTGGGGTCTTATGCAGAGCTTTATATAACACCTGTAAAAGATTTAGAATGG

TyrPheGluMetAspValAsnAsnSerAspSerAspSerThrGlyIleProValSerPhe
5760 TATTTTGAAATGGATGTTAATAATAGTGATTCAGATTCTACAGGTATACCTGTTAGTTTT

AlaSerThrThrGlyIleThrTrpTyrLeuProGluPheOC
5820 GCTTCTACTACAGGATAAAGTTGGTATTTGCCAGAATTTTAATTATAAA6AAAAATTTTAT

5880 ATGACAAAATAAAAAATAAGAGGTATTTATTTTTTAATAATAAATACCTCTCTTATTTT

5940 ATTTATA8TTTTTTAATTTAATCTCTTCTCTTTGTGCTAACATTAAAAGTCTTAAAAGT

-50-

6000 GTAAGTATAGCAGTAACAGCAGCAGCAACATAAGTCAAAGCAGCAGCAGATAAAACTTTT

6060 TTAGCTCCGTCAAGTTCTTCACTGTCTAAAAATCCGCCTTTATCCAATATTTTTATAGCT

Appendix 2 B

Figure # --DNA Sequence of B204 Genes Encoding 39 kDa Antigens 5-8

*5' end of genomic DNA

1 CTGTATCTGTTTATTATCAGTCAGCATTTCATTCTTATTATTATGGCGATAAAAAATATAA

61 AAATGGAACTATAGTCGGAACTTAATAGGAAGCGT4GCCAAAGAAAATAAAGATGATA

121 TACCAAAATTAAAGAACTATTTTAATAATTCTGTAAAAATAAAAGCTGAGAAATTTGGTA

181 AAAAATGGAAAGACTATTATGAAAGCAGAAATCTTATAAATAATTTTAATTTATAATATA

241 TTACCGATAAAAACAGCATTAAATGTTTTTTATAAATACTGTATTTACAAAAATATAGCAT

301 TATGTATTAAATAAACTTTAATATATCTTTTATATTTTATACTTTTGACTGTAATATTAA

361 TTATCATTTTATATAACTATACCTAAATAGATATATTACTATAATTGTATCACAAATTAGA

421 TATAAACCATCATATTTAACATAAGAAAAATATATTATTATAAAATATAAAAAATTTT

481 GTATTTTTTATAGAATTTTAA9AAAAATTTTATATAATAATATTCATATATATTAGGA4

* Presumed translational start of Copy 8

MetLys PheLeuLeuThrValLeuAlaIleLeuThrIleAlaSerGly

541 AAAATAAAAATGAAAAA6TTTTTATTAAACAGTGCCTGGCTATTTTAAACAATAGCTAGCGGA

SerValPheGlyMetTyrGlyAlaAspAsnThrTrpLeuPhePheLeuIleHisGlyAsn

601 TCAGTGTTTGGTATGTATGGTGCAGACAACACATGGCTTTTCTTCATACATGGCAAC

GlnPheArgAlaArgMetAsnGlnLeuGlyPheThrLeuGlyAsnGlyIleIleLysGly

661 CAATTCAGAGCTAGAATGAACCAATTAGGTTTCACTCTAGGTAACGGCATTATTAAAGGT

ThrPheGlyPheLysAlaAsnThrLeuIleAsnGlySerIleLeuAsnThrGlyAsnLys

721 ACTTTTCGGTTTCAAAGCTAATACGCTTATTAAACGGAAGCATCTTAAATACAGGCAATAAA

GluAsnGlnAsnProLeuGluAlaThrIleSerAlaGlyIleGlyTyrThrGlyAspGly

781 GAAAACCAAATCCATTAGAAGCTACTATTTCTGCTGGTATAGGTTACACAGGTGATGGT

PheGlyValGlyValGlyTyrAsnTyrThrTyrThrAlaAlaAsnThrIleGlnThrLys

841 TTTGGTGTGGTGTGGTTATACTATACTTATACTGCTGCAAACTACTATTCAAACCAA

AlaAlaLysGlyIleAsnThrHisThrProValIleThrPheAsnAlaValAsnAsnAsn

901 GCTGCTAAAGGAATAAATACTCATACACCTGTTATTACATTCAATGCTGTTAATAACAAT

LeuArgValAlaIleProValSerIleAlaValGluLysAspIleGlyLysLeuGlyAsn

961 TTAAGAGTAGCTATACCTGTAAGTATAGCTGTAGAAAAGATATAGGTAAATTAGGTAAT

MetAspArgLysAspTyrLeuGlyLeuSerIleProAlaGlnIleArgTyrTyrThrGly

1021 ATGGATAGAAAAGATTATTTAGGTTTAAAGCATACCTGCTCAAATAAGATATTATACAGGA

IleAspAlaPheAsnTyrIleArgPheGluPheAsnTyrGlyLeuAsnLysTyrAsnGly

1081 ATAGATGCTTTCAACTATATAAGATTGAATTCAATTACGGATTAAATAAATAATATGGT

ValGluAsnAsnThrThrThrGluTyrGlnAlaGlnThrIleSerPheGlnLeuArgLeu

1141 GTTGAAAACAATACAACCTACAGAATATCAAGCACAGACTATAAGCTTCCAATTAAGACTT

-56-

1201 HisPheLeuAsnThrValLeuGlyAsnAsnValThrValAsnProPheLeuArgValAsp
CATTTCTTAAATACAGTTCTTTGGAAATAATGTAACGTAAACCCATTCTTAAGAGTTGAT

1261 PheAlaSerThrValGlyAlaLysGlyLysGlyAsnValValPheProAlaAlaThrAla
TTTGCTTCTACTGTAGGTGCTAAAGGAAAAGGAAATGTAGTATTCCCAGCAGCTACAGCT

1321 PheAspGlyArgLeuThrAlaTrp AlaAsnAlaTrpAlaAspAspProHisSerIle
TTTGATGGAAGACTTACAGCTTGGGCTGCAAATGCTTGGGCAGATGATCCACACAGCATT

1381 TyrAspArgGluLeuTyrAspLeuLysIleIleProSerValSerLeuSerValAsnThr
TATGACAGAGAACTTTATGACTTGAAAATCATACCTAGCGTATCTTTAAGTGTTAATACA

1441 AspTyrValAsnLeuIlePheGluProGlyLeuGlyTyrArgValGlnAspAspGlyVal
GACTATGTTAATTTAATATTTGAACCTGGTTTAGGATACAGAGTACAAGATGATGGTGTA

1501 LysGlySerLysLeuThrHisThrLeuTyrTrpGlnAlaTyrGlyGluIleTyrIleArg
AAAGGAAGCAAACCTTACTCATACATTATACTGGCAGGCTTACGGAGAAATATATATCAGA

1561 ProValGlnAspLeuGluTrpTyrPheGluMetAspValAsnAsnGlyValProLysLeu
CCTGTTTCAGGATCTTGAATGGTATTTGAAATGGATGTTAATAACGGAGTACCTAAACTT

1621 GlnGlyAsnProIleAlaSerGlyAsnSerMetProValValPheGlyAlaAsnThrGly
CAAGGAAATCCTATTGCTTCAGGAAATTCATGCCTGTTGTATTTCGGAGCTAATACTGGT

1681 IleThrTrpTyrLeuProAlaLeuGln
ATAACTTGGTATTTACCTGCTTTACAATAATAATAACAGTGAATAATCAGAAATAACTG

1741 ATTTTAAATACAAGGGGAGCTTCXTTAAAAAGCTCCTCTTTTTTATGAGAAAAACAAATA

1801 TAAAGCTTNNGTTCATAATTATCAATTTAATCAAAACATACTAAGCAAATTCAAAATTTT

1861 ATAAAAAATAATTATTCAATAAAAATTCATAATTTTTTATTAAAAAATATGYAAATATG

1921 TATTTGAATGACTATTTTTTTATATAATTCCTTTTATATGTATTGACATTTTATACTTT

1981 TTCCTATAATGAAAGTAGAGGATCTTTAGGTGTCTATTGAGTTTTTAAGCGATCTCTCAA

* Presumed translational
MetLysLysValLeuLeu.

2041 TAGACACTGTCAAAAGATTATAAAAAAATTAGGAGAAAAACAATGAAAAAGTTTTATTG

start of Copy 6

2101 ThrAlaMetAlaLeuLeuThrIleAlaSerAlaSerAlaPheGlyMetTyrGlyAspArg
ACAGCTATGGCATTATTGACTATAGCTAGTGCATCTGCTTTCGGTATGTATGGCGACAGA

2161 AspSerTrpIleAspPheLeuThrHisGlyAsnGlnPheArgAlaArgMetAspGlnLeu
GATTCTTGGATCGACTTCCTTACTCATGGTAATCAGTTCAGAGCTAGAATGGATCAATTA

2221 GlyPheValLeuGlyAsnGlyThrIleLysGlyThrPheGlyPheArgSerGlnAlaIle
GGTTTCGTTTTAGGTAACGGTACTATTAAAGGTACTTTCGGTTTTAGATCTCAAGCTATT

2281 GlyThrAlaLeuGlyAsnIleIleSerGlyAsnThrGlyAsnValAspLeuGlnThrThr
GGAACAGCATTAGGTAATATCATTTTCAGGTAATACTGGAAATGTAGATTACAACTACT

IleSerAlaGlyIleGlyTyrThrSerGluProPheGlyIleGlyValGlyTyrAsnTyr
2341 ATTTCTGCTGGTATAGGTTATACTTCTGAGCCTTTCGGTATTGGCGTAGGTTATAACTAC

ThrTyrValAsnProArgLeuGlyValHisThrProValLeuMetIleAsnAlaLeuAsn
2401 ACTTATGTAAATCCTAGATTAGGCGTTCATACTCCTGTACTTATGATCAATGCTTTAAAC

AsnAsnLeuArgIleAlaValProValGlnIleAlaValSerHisAspProPheAsnAsp
2461 AACAACTTAAGAATAGCAGTTCCTGTTCAAATAGCTGTAAGTCATGATCCTTTCAATGAT

SerAlaLysPheProTyrSerSerSerThrLysAspTyrMetGlyIleSerThrAspIle
2521 TCTGCTAAATTCCTTATTCATCATCTACAAAAGATTATATGGGTATAAGCACTGATATA

GlnLeuArgTyrTyrThrGlyIleAspAlaPheAsnAlaIleArgLeuTyrPheLysTyr
2581 CAATTAAGATACTATACTGGTATAGATGCTTTCAATGCTATAAGATTATACTTCAAATAC

GlyGlnAlaGlyPheLysThrAlaAsnGlyAlaGlyAlaSerGluTyrPheAlaGlnSer
2641 GGACAAGCTGGATTTAAACAGCTAACGGAGCTGGAGCTAGTGAGTATTTTGCTCAGTCA

LeuGlyPheGluAlaArgPheTyrPheLeuAsnThrProValGlyAsnValThrIleAsn
2701 TTAGGTTTTGAAGCTAGATTCTATTTCTTGAATACTCCTGTTGGAAACGTAACATCAAT

ProPheIleLysValValTyrAsnThrAlaLeuLysGlyValSerArgThrValArgAla
2761 CCTTTCATCAAAGTTGTTTATAACACAGCTTTAAAGGTGTAAGCAGAACTGTAAGAGCT

GlyGluAlaValGlnAsnThrValSerGlyTyrHisProSerAsnProAsnTyrLysLeu
2821 GGAGAAGCTGTACAAAATACTGTTTCTGGTTATCATCCTTCTAATCCTAATTATAAATTA

AspAlaPheAlaGlyArgTyrIleGlyLysAspPheLysTrpAspSerAsnProTyrAsp
2881 GATGCATTTGCTGGTAGATACATTGGTAAAGATTTCAAATGGGATTCAAATCCTTATGAT

ValLysAlaGlnAlaValLeuGlyIleThrAlaAsnSerAspValValSerLeuTyrVal
2941 GTAAAGCTCAGGCTGTATTAGGTATCACTGCTAACAGCGATGTAGTATCTCTTTATGTT

GluProSerLeuGlyTyrGlnAlaThrTyrLeuGlyLysAsnIleSerGluAsnProTyr
3001 GAGCCTTCTTTAGGTTATCAAGCTACATATTTAGGAAAAACATATCTGAAAATCCATAT

LeuAsnIleAspSerLysValGlnHisSerLeuAlaTrpGlyAlaTyrAlaGluLeuTyr
3061 TTAAATATAGATTCTAAAGTACAACATAGCTTAGCTTGGGGTGCTTATGCAGAACTTTAT

ValArgProValGlnAspLeuGluTrpTyrPheGluMetAspValAsnAsnGlyGlyThr
3121 GTAAGACCTGTTCAAGATCTTGAATGGTACTTCGAGATGGATGTTAATAATGGCGGTACA

ArgGlnGluSerGlyIleProValTyrPheLysSerThrThrGlyIleThrTrpTyrLeu
3181 AGACAAGAATCTGGTATCCCTGTATACTTTAAATCTACTACAGGTATAACTTGGTATTTA

ProAlaPheAsn
3241 CCTGCTTTCAATTAATTAGAAGTTAATTAATAGAAATTAATGAGGCTGGCCTTTAATAGG

TTGGCCTCTTTTTTATTAATTTTCATATTGCAAATAGTTTATTCATTATTATATAAGTT
3301

TACATTTTATTGTTTTGCAATATAATATTTTTATCTATAATCTACCTATATAAATTATAT
3361

-58-

* Presumed translational
MetLysLysIleMetLeuAla

3421 AATAGAAGAAAATATTTTATGTTTAGGAGAACAGATAAAATGAAAAAATTATGCTGGCA

start of Copy 7

AlaIleAlaIleLeuThrIlePheSerAlaSerAlaLeuGlyMetTyrGlyAspGlnAsp

3481 GCTATTGCTATATTAAGTATATTTAGTGCATGTGCTTTAGGAATGTATGGAGATCAAGAT

AspTrpIleAspPheLeuThrAspGlyAsnGlnLeuArgAlaArgMetAspGlnLeuGly

3541 GACTGGATTGATTTTCTTACAGATGGAAATCAATTAAGAGCCAGAATGGATCAATTAGGA

PheValLeuGlyAsnAsnThrIleLysGlyThrPheGlyLeuArgThrGlnAspAlaVal

3601 TTTGTACTTGGAAACAATACTATTAAAGGTACTTTCGGACTTAGAACTCAAGATGCCGTA

ThrSerLeuGlySerIleIleSerGlyLysThrAspAsnLeuGlyLeuAspAlaThrVal

3661 ACATCATTGGGAAGTATAATTCAGGTAAACAGATAATTTAGGATTAGATGCTACTGTT

SerMetGlyIleGlyTyrThrSerAspIlePheGlyIleGlyLeuGlyTyrAsnPheThr

3721 TCTATGGGAATAGGATACACTTCTGATATTTTCGGCATTGGCTTAGGATATAATTTTACA

TyrTyrAsnSerThrLeuGlyValHisThrProValLeuMetValAsnAlaLeuAsnAsn

3781 TATTATAACAGCACTTTAGCGCTTCATACTCCTGTACTTATGGTCAATGCTTTAAATAAT

AsnLeuArgIleAlaIleProIleGlnIleAlaAlaSerLysAspProPheGlyLysTyr

3841 AATTTAAGAATAGCAATACCTATACAAATAGCTGCATCAAAGATCCTTTCGGAAAATAT

ThrIleSerGlnTyrLysAspTyrLeuGlyIleSerThrAspIleGlnIleArgTyrTyr

3901 ACTATCAGTCAATATAAAGACTATTTAGGAATAAGCACAGATATACAAATAAGATACTAT

ThrGluIleAspValPheAsnGlnValArgLeuTyrIleLysTyrGlyGlnSerGlyTyr

3961 ACAGAAATAGATGTATTCAATCAAGTAAGATTATACATCAAATATGGACAGTCAGGTTAT

LysAsnValLysAsnAsnPheAspMetPheAlaGlnSerPheGlyPheGluThrArgLeu

4021 AAAAATGTTAAAAATAATTTTGATATGTTTGCTCAATCATTGGTTTTGAACTAGACTA

TyrPheLeuAsnArgThrIleGlyAsnValAsnIleAsnProPheIleLysValSerTyr

4081 TATTTCTTAACCGCACAAATTGGAAATGTAAATATTAATCCTTTTATAAAGTTTCATAT

AsnThrAlaLeuAlaSerSerAspValMetValArgAlaGlyGluSerLeuValAsnThr

4141 AATACAGCTTTAGCCAGCAGTGATGTAATGGTTAGAGCAGGAGAATCTCTGTAAATACT

ThrTyrSerLysLysGluAsnLysTrpGluLysAsnProTyrAsnValThrAlaAlaAla

4201 ACTTATAGTAAAAAGAAAATAAATGGGAAAAAAATCCTTATAATGTAAGTCTGCTGCT

ValLeuGlyLeuThrAlaAsnSerAspMetLeuSerLeuCysValGluProSerLeuGly

4261 GTATTAGGATTAAGTCTAACAGTGATATGCTATCTCTTTGTGTGAACCTTCTTTAGGA

TyrAsnAlaValTyrLysGlyLysTyrLysThrAspSerLysTyrTyrLysValGlnHis

4321 TACAATGCCGTTTATAAAGGAAAATATAAACTGATAGTAAATATTATAAAGTACAGCAT

AsnLeuTyrTrpGlyAlaTyrAlaGluLeuTyrIleThrProValGlnAspIleGluTrp

4381 AATTTATATTGGGGAGCTTATGCAGAACTTTATATTACTCCCGTTCAAGATATTGAATGG

-59-

4441 TyrPheGluMetAspIleAsnAsnGlyAsnSerArgGlnThrSerSerIleProIleTyr
TATTTTGAAATGGACATTAATAATGGTAATTCAAGACAGACTTCTTCTATACCTATATAC

4501 PheGluSerThrThrGlyIleThrTrpTyrLeuProGluLeu
TTTGAATCTACTACAGGGATAACTTGGTATTTGCCTGAATTATAATAAAAAATATTATTTT

4561 GTATAATTTAAACATTTATTGAATAAAAAAATAAAAAATTAATAATTCTATAAAAAATT

4621 TAAATTAATACTTTACATATTTAAATAAATAGCTGTTTTTTATTAAATCTAAGTATTATTT

4681 TAAAAAATAATAATTTTATTTATATTAGTGTTGACTTTTTTATGTAAATTCATTAGAATA

4741 AAAGTGTTGGTAAGATATAGACATTAAAAAAGATTGCTTATAATGTCTATAACCATATT

4801 AGGAACAATAAAGATCCTCTAAATCCTTATTATTCCAATTGACACATTAGTAAAAACAGC

4861 ATAATTTTTCCATTTAAATATCTGTTTTTACGGAGAATAAAGATCCTCTAAATCTTATTA

4921 TTCTCACTGATTTTTTATAATCCGTTAGAATTATATAGATGTTCTAACTCATCTGGAAAT

4981 AACAAAGATCCTCTGTCTTTGTTGTTTCTAAAATTATTTTGGAGTTACTACTTACAATG

*Presumed translational start
MetLysLysValLeuLeuThrAlaMet

5041 AGTATTAECTATATAAAAATTTTAGGAGAAAAATAATGAAAAAGTTTTATTGACAGCTAT

* Amino terminus of Copy 5

5101 of Copy 5
AlaLeuLeuThrIleAlaSerAlaSerAlaPheGlyMetTyrGlyAspArgAspSerTrp
GGCATTATTGACTATAGCTAGTGCATCTGCTTTCGGTATGTATGGCGACAGAGATTCTTG

5161 IleAspPheLeuThrHisGlyAsnGlnPheArgAlaArgMetAspGlnLeuGlyPheVal
GATCGACTTCCTTACTCATGGTAATCAGTTCAGAGCTAGAATGGATCAATTAGGTTTCGT

5221 LeuGlyAsnGlyThrIleLysGlyThrPheGlyPheArgSerGlnAlaIleGlyThrAla
TTTAGGTAACGGTACTATTAAAGGTACTTTCGGTTTTAGATCTCAAGCTATTGGAACAGC

5281 LeuGlyAsnIleIleSerGlyAsnThrGlyAsnValAspLeuGlnThrThrIleSerAla
ATTAGGTAATATCATTTTCAGGTAATACTGGAAATGTAGATTTACAACTACTATTTCTGC

5341 GlyIleGlyTyrThrSerGluProPheGlyIleGlyValGlyTyrAsnTyrThrTyrVal
TGGTATAGGTTATACTTCTGAGCCTTTCGGTATTGGCGTAGGTTATAACTACACTTATGT

5401 AsnProArgLeuGlyValHisThrProValLeuMetIleAsnAlaLeuAsnAsnAsnLeu
AAATCCTAGATTAGGCGTTCATACTCCTGTACTTATGATCAATGCTTTAAACAACAACCTT

5461 ArgIleAlaValProValGlnIleAlaValSerHisAspProPheAsnAspSerAlaLys
AAGAATAGCAGTTCCTGTTCAAATAGCTGTAAGTCATGATCCTTTCAATGATTCTGCTAA

5521 PheProTyrSerSerSerThrLysAspTyrMetGlyIleSerThrAspIleGlnLeuArg
ATTCCCTTATTCATCATCTACAAAAGATTATATGGGTATAAGCACTGATATACAATTAAG

5581 TyrTyrThrGlyIleAspAlaPheAsnAlaIleArgLeuTyrPheLysTyrGlyGlnAla
ATACTATACTGGTATAGATGCTTTCAATGCTATAAGATTATACTTCAAATACGGACAAGC

5641 GlyPheLysThrAlaAsnGlyAlaSerGluTyrPheAlaGlnSerLeuGlyPheGluAla
TGGATTTAAACAGCTAACGGAGCTAGTGAGTATTTTGCTCAGTCATTAGGTTTTGAAGC

5701 ArgPheTyrPheLeuAsnThrProValGlyAsnValThrIleAsnProPheIleLysVal
TAGATTCTATTTCTTGAATACTCCTGTTGGAAACGTAACATCAATCCTTTCATCAAAGT

5761 ValTyrAsnThrAlaLeuLysGlyValSerArgThrValArgAlaGlyGluAlaValGln
TGTTTATAACACAGCTTTAAAAGGTGTAAGCAGAACTGTAAGAGCTGGAGAAGCTGTACA

5821 AsnThrValSerGlyTyrAsnProTyrAspProAsnTyrLysLeuAspAlaPheAlaGly
AAATACTGTTTCTGGTTATAATCCTTATGATCCTAATTATAAATTAGATGCATTTGCTGG

5881 ArgTyrIleGlyLysAspPheLysTrpAspSerAsnProTyrAspValLysAlaGlnAla
TAGATACATTGGTAAAGATTCAAATGGGATTCAAATCCTTATGATGTAAAAGCTCAGGC

5941 ValLeuGlyIleThrAlaAsnSerAspValValSerLeuTyrValGluProSerLeuGly
TGTATTAGGTATCACTGCTAACAGCGATGTAGTATCTCTTTATGTTGAGCCTCTTTAGG

6001 TyrGlnAlaThrTyrLeuGlyLysHisIleSerGluAsnProTyrLeuAsnIleAspSer
TTATCAAGCTACATATTTAGGAAAACACATATCTGAAAATCCATATTTAAATATAGATTC

6061 LysValGlnHisSerLeuAlaTrpGlyAlaTyrAlaGluLeuTyrValArgProValGln
TAAAGTACAACATAGCTTAGCTTGGGGTGCTTATGCAGAACTTTATGTAAGACCTGTTCA

6121 AspLeuGluTrpTyrPheGluMetAspIleAsnAsnSerAspSerLysArgAsnGlyVal
AGATCTTGAATGGTACTTCGAGATGGACATCAATAACTCTGATTCAAAAAGAAATGGTGT

6181 ProValAsnPheAlaThrSerThrGlyIleThrTrpTyrLeuProAlaLeuGlyGlyAla
TCCTGTAACTTCGCAACTTCTACAGGTATAACTTGGTACTTACCTGCTTTAGGCGGTGC

6241 Gln
TCAATAATTAATTTCTGTAAATTAAAGAATTTACAGAGGCTGGCTTTAATAAAAAAGTCA

6301 GTCTCTTTTTTTATCGCATATTTTCATATAATTAAACAAAAATATTTACATTATATAATT

-51-

Appendix 3

Nucleotide sequence of T. hyo gene insert of pTrep106

(1) (in) (er) 1--T. hyo DNA --

1 GAATTCC CTATTTTATATTTTCTAGTATAATAAAAAATATAAACTTCAATTTAGGTATGTATA
CTTAAGG GATAAAATATAAAATCATATTATTTTATATTTGAAGTTAAATCCATACATAT

1 ECOR1

(Signal peptide sequence --)

63 MetLysLysPhePheLeuIleMetThrValLeuLeuSerMetSerTyrCysSerIlePhe
ATGAAAAAGTTTTTCTAATTATGACAGTATTATTAAGTATGTCAATTTGTTCAATTTTT
TACTTTTTCAAAAAGATTAACTGTCTATAATAATTTCATACAGTATAACAAGTTAAAAA

1--begin mature protein -->

123 GlyMetTyrGlyAspGlnAspAspTropIleAspPheLeuThrAspGlyAsnGlnPheArg
GGTATGTATGGAGATCAGGACGATTGGATTGATTTTCTTACAGACGGCAATCAGTTTAGA
CCATACATACCTCTAGTCCTGCTAACCTAACTAAAAGAATGTCTGCCGTTAGTCAATCT

183 AlaArgMetAspGlnLeuGlyPheValLeuGlyAsnSerThrIleLysGlyThrPheGly
GCTAGAATGGATCAATTAGGATTTGTTTTAGSTAATAGTACTATTAAAGGTACTTTCGGT
CGATCTTACCTAGTTAATCCTAAACAAAATCCATTATCATGATAATTTCCATGAAAGCCA

243 PheArgThrGlnSerSerSerThrGlnLeuGlyTyrIleLeuLeuAsnAsnAsnLeuGly
TTTAGAACTCAAAGTTCATCAACTCAATTAGGATATATTTTGTGAAATAATAATCTTGGT
AAATCTTGAGTTTCAAGTAGTTGAGTTAATCCTATATAAAACAACCTATTATTAGAACCA

303 ThrTyrLeuGlyAlaThrIleSerGlyGlyIleGlyTyrThrSerGluAlaPheSerIle
ACTTATTTGGGAGCAACTATTTCTGGCGGTATAGGATATACTTCTGAGGCTTTTAGTATA
TGAATAAACCCCTCGTTGATAAAGACCGCCATATCCTATATGAAGACTCCGAAAATCATAT

363 GlyIleGlyTyrAsnTyrThrSerHisSerLeuPheProThrSerAspAsnPheGlySer
GGCATAGGCTATAATTATACCAGCCATTCCTTATTTCTTACTAGCGATAACTTTGGTTCT
CCGTATCCGATATTAATATGTCGSTAAGGAATAAAGGATGATCGCTATTGAAACCAAGA

423 HisThrProValLeuMetIleAsnAlaLeuAsnAspAsnLeuArgIleValIleProVal
CATACTCCAGTACTTATGATTAATGCTTTAAATGATAATTTGAGGATAGTTATTCCTGTG
GTATGAGGTCATGAATACTAATTACGAAATTTACTATTAACTCCTATCAATAAGBACAC

483 GlnIleLeuValHisAsnGluSerIleAspGlnLeuGlyTyrTyrArgAspAsnTyrLeu
CAAATATTATACATAATGAAAGTATTGATCAACTTGGTACTATAGAGATAATTATTTA
GTTTATAATCATGTATTACTTTCACTAGTTGAACCAATGATATCTCTATTAATAAAT

543 GlyIleSerThrAspThrGlnIleArgTyrTyrThrGlyIleAspAlaPheAsnGluIle
GGTATAAGTACTGATACGCAATAAGATATTATACAGGCATAGATGCTTTTAAATGAAATA
CCATATTCATGACTATGCTTTTATTCTATAATATGTCGATCTACGAAAATTACTTTAT

603 ArgLeuTyrValLysTyrGlyGlnLeuGlyTyrLysIleAsnProHisAspThrIleAsn
AGATTATATGTAAATATGGGCAATTAGGATATAAAATTAATCCGCATGATACTATAAAT
TCTAATATACATTTTATACCCGTTAATCCTATATTTTAATTAGGCGTACTATGATATTTA

663 TyrThrGlnGluValLeuAlaArgSerPheGlyPheGluThrArgPheTyrPheLeuAsn
TATACACAAGAAGTTTTAGCAAGATCATTTGGTTTTGAAACAAGATTCTATTTTTTGAAT
ATATGTGTTCTTCAAAATCGTTCTAGTAAACCAAACTTTGTTCTAAGATAAAAACTTA

-62-

723 ThrAlaValGlyAsnValThrIleAsnProPheIleLysValAlaTyrAsnThrAlaLeu
ACTGCTGTTGAAATGTAACATCAATCCTTTTATTAAAGTAGCATATAATACAGCTTTG
TGACGACAACTTTACATTGATAGTTAGGAAAATAATTTTCATCGTATATTATGTCGAAAC

783 HisGlyTyrSerThrMetValArgAlaLeuAspGlyMetTyrGluGluIleGluGlyTyr
CATGGATATAGTACCATGGTAAGAGCATTGGATGGTATGTATGAAGAAATAGAAGGTTAT
GTACCTATATCATGGTACCATTCTCGTAACCTACCATACATACTTCTTTATCTTCCAATA

843 TyrProAspSerProAlaGlnSerTyrGluAspIleAsnValLysTrpAspLysAsnPro
TATCCAGATAGTCTCTGCTCAATCATATGAAGATATTAATGTAAATGGGATAAGAATCCT
ATAGGTCTATCAGGACGAGTTAGTATACTTCTATAATTACAATTTACCCTATTCTTAGGA

903 TyrAspValThrValGlnAlaValLeuGlyValThrAlaAsnSerAspIleValSerLeu
TATGATGTAACGTGTCAGGCAGTATTGGGAGTAACTGCTAATAGCGATATAGTATCACTT
ATACTACATTGACACGTCCGTCTATAACCTTCATTGACGATTATCGTATATCATAGTGAA

963 TyrValGluProSerLeuGlyTyrArgAlaLysTyrLeuGlyLysLeuThrTyrGluAsp
TATGTTGAGCCTTCTTTAGGTTATAGGGCTAAATATTTAGGAAAATTAACATATGAAGAT
ATACAACTCGGAAGAAATCCAATATCCCGATTTATAAATCCTTTTAATTGTATACTTCTA

1023 ProAspGlyLysValAsnPheAspPheLysValAsnHisTyrLeuSerTrpGlyAlaTyr
CCAGATGGAAAAGTTAATTTTGATTTTAAAGTTAATCATTATTTATCTTGGGGTGCTTAT
GGTCTACCTTTTCAATTAAACTAAAATTTCAATTAGTAATAAATAGAACCCACGAATA

1083 AlaGluLeuTyrIleThrProValLysAspLeuGluTrpTyrPheGluMetAspValAsn
GCAGAGCTTTATATAACACCGGTAAAAGATTTAGAATGGTATTTTGAAATGGATGTTAAT
CGTCTCGAAATATATTGTGGCCATTTTCTAAATCTTACCATAAACTTTACCTACAATTA

1143 AsnSerAspSerAspSerThrGlyIleProValSerPheAlaSerThrThrGlyIleThr
AATAGTGATTCAGATTCTACAGGTATACCTGTTAGTTTTGCTTCTACTACAGGAATAACT
TTATCACTAAGTCTAAGATGTCCATATGGACAATCAAAACGAAGATGATGTCTTATTGA

1203 TrpTyrLeuProGluPheOC
TGGTATTTGCCABAATTTTAATTATAAAGCAAATTTTATATGATAAAATAAAAAATGTGG
ACCATAAACGGTCTTAAATTAATATTTCTTTTAAATATACTATTTTATTTTTTACACC

1263 GGTATTTATTATTAAAAAATAAATACCCACATTTTATTAAATAATTTTTTCAGTAATTT
CCATAAATAATAATTTTTTATTTATGGGTGTAAATAATTTATTAAAAAAGTCATTAA

(begin gene encoding a related protein-->

1323 TACATTTATATATTTTTTATATAATAAAAAATATAAACTTAAATTTAGGTATATACAATG Met
ATGTAAATATATAAAAAATCATATTATTTTATATTTGAATTTAAATCCATATATGTTAC

1383 LysLysIlePheLeuIleMetThrValLeuLeuSerMetSerTyrCysSerIlePheGly
AAAAAAATTTTTCTAATTATGACAGTATTATTAAGTATGTCATATTGTTCAATATTTGGT
TTTTTTTAAAAAGATTAATACTGTCTATAATAATTCATACAGTATAACAAGTTATAAACCA

(linker)

1443 MetTyrGlyAspGlnAspAspTrpIleAspPheLeuIleAspGlyAsnGlnPheArg
ATGTATGGAGATCAGGACGATTGGATTGATTTTCTTATAGACGGTAATCAGTTTATAG GGAATTC
TACATACCTCTAGTCTCTGCTAACCTAACTAAAAGAAATATCTGCCATTAGTCAAACTC CTTAAG

1502 ECOR1

-63-

Appendix 4

Partial DNA sequence of expression plasmid pTrep301

```

(plasmid pUC9---) 1/--- E. coli lacZ ----- 1 -T. hyo
MetThrMetIleThrProSerLeuHisAlaSerTyrCys
1 AATTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCCTCATATTGT
TTAAAGTGTGTCCTTTGTCGATACTGGTACTAATGCGGTTTCAACGTACGGAGTATAACA
39 HINDIII.

signal ---->1<---mature T. hyo 39 kDa protein---
SerIlePheGlyMetTyrGlyAspGlnAspAspTrpIleAspPheLeuThrAspGlyAsn
61 TCAATTTTTGGTATGTATGGAGATCAGGACGATTGGATTGATTTTCTTACAGACGGCAAT
AGTTAAAAACCATACATACCTCTAGTCCTGCTAACCTAACTAAAAGAATGTCTGCCGTTA

GlnPheArgAlaArgMetAspGlnLeuGlyPheValLeuGlyAsnSerThrIleLysGly...
121 CAGTTTAGAGCTAGAATGGATCAATTAGGATTTGTTTTAGGTAATAGTACTATTAAAGGT...
GTCAAATCTCGATCTTACCTAGTTAATCCTAAACAAAATCCATTATCATGATAATTTCCA...

```

The first 25 amino acids of this recombinant fusion protein were determined on protein isolated from *E. coli* transformed with plasmid pTrep301. The protein was isolated as a Triton X-100 insoluble aggregate and was further purified by preparative SDS gel electrophoresis.

Appendix

--Predicted amino acid sequence from PCR derived T_hyo
(B204) clones

pTrep605--Copy2

10	20	30	40	50	60
MTMITNRGSM	YGDQDDWIDF	LTDGNQFRAR	MDQLGFVLGN	STIKGTFGFR	SQSLSTQLGY
70	80	90	100	110	120
ILAIYKDYTY	LGATISGGIG	YTSEAFSIGL	GYNYYTTPLP	SYNFGSHTPV	LMINALNDNL
130	140	150	160	170	180
RIVIPVQILV	HDGNMNMMDN	INYLYNFI	STDQIRYYT	GIDAFNEIRL	YVKYQQLGYK
190	200	210	220	230	240
GGSYTDKSYD	EEFFARSFGF	ETRFYFLNTA	VGNVTINPFI	KVAYNTALHG	FSTMVRSLS
250	260	270	280	290	300
VIEEIEGYSS	DRTAKAAGNI	NAKWDKNPYD	VTVQAVLGVT	ANSDIVSLYV	EPSLGYRAKY
310	320	330	340	350	360
LGKLTIEDPD	GKVNFDKVN	HYLSWCAYAE	LYITPVKDL	WYFEMDVNNS	DSDSTGIPVS
370					
FASTTGITWY	LPEF				

pTrep604--Copy3

10	20	30	40	50	60
MTMITNRGSM	YGDQDDWIDF	LTDGNQFRAR	MDQFGFVLGN	STIKGTFGFR	SQSLSTQLGY
70	80	90	100	110	120
ILAIYKDYTY	LGATISGGIG	YTSEAFSIGL	GYNYYTTPLP	SDNFGSHTPV	LMINALNDNL
130	140	150	160	170	180
RIVIPVQILV	YNGNVQKVDK	QGNISYSHDY	LGISTDTQIR	YYTGIDAFNE	IRLYVKYQQL
190	200	210	220	230	240
GYKNAPYVGK	NYEEFFSRS	FGFETRFYFL	NTAVGNVTIN	PFIKVAYNTA	LHGFSTMIRA
250	260	270	280	290	300
LDSMFEP	IEG YSSDRPVSSQ	ANINAKWDKN	PYDVTVQAVL	GVTANSDIVS	LYVEPSLGYR
310	320	330	340	350	360
AKYLKGLTYE	DPDGKVNLD	KVNHYLSWGA	YAELYITPVK	DLEWYFEMDV	NNSDSDSTGI
370	377				
PVSEFASTGI	TWYLPEF				

-65-

pTrep345--Copy4

10	20	30	40	50	60
MTMITNRGSM	YGDQDDWIDE	LTDGNQFRAR	MDQFGFVLGN	NTIKGTFGFR	SQSLSTHLGY
70	80	90	100	110	120
ILLNNNEGTY	FGTTISCGIG	YTSEAFSIGI	GYNYTTPLPI	SDNEGSHTPV	LMINALNDNL
130	140	150	160	170	180
RIVIPVQILV	YNGNIQKVDK	QGNIHDTYDY	LGISTDTQIR	YYTGIDAFNE	IRLYIKYGQL
190	200	210	220	230	240
GYKNAPYVGK	NYEEELFSRS	FGFETRFYFL	NTTVGNVTIN	PFIKVAYNTA	LHGVGTMIRA
250	260	270	280	290	300
LDTMLQPIED	YYPDRPVSSQ	VDIDYKLDKN	PYDVTVQAVL	GVTANSDIVS	LYVEPSLGYK
310	320	330	340	350	360
AKYLGKMQDE	KVNLDKFNH	YLSWGAYAEL	YITPVKDLEW	YFEMDVNNSD	SDSTGIPVSF
370					
ASTTGITWYL	PEF				

pTrep613--Copy5

10	20	30	40	50	60
MTMITNRGSM	YGDRDSWIDE	LTHGNQFRAR	MDQLGFVLGN	GTIKGTFGFR	SQAIGTALGN
70	80	90	100	110	120
IISGNTGNVD	LQTTISAGIG	YTSEPFGIGV	GYNYTYVNPR	LGVHTPVLM	NALNNNLRIA
130	140	150	160	170	180
VPVQIAVSHD	PFNDSAKFPY	SSSTKDYMGI	STDIQLRYT	GIDAFNAIRL	YFKYQGAGFK
190	200	210	220	230	240
TANGASEYFA	QSLGFEARFY	FLNTPVGNVT	INPFIKVVYN	TALKGVSRTV	RAGEAVQNTV
250	260	270	280	290	300
SGYNPYDPNY	KLDAFAGRYI	GKDFKWDSNP	YDVKAQAVLG	ITANSVVVSL	YVEPSLGYQA
310	320	330	340	350	360
TYLGKHISEN	PYLNIDSKVQ	HSLANGAYAE	LYVRPVQDLE	WYFEMDINNS	DSKRNGVPVN
370	378				
FATSTGITWY	LPALGGAQ				

-66-

pTrep620--Copy7

10	20	30	40	50	60
MTMITNRGSM	YGDQDDWIDF	LTDGNQLRAR	MDQLGFVLGN	NTIKGTFLGR	TQDAVTSLGS
70	80	90	100	110	120
IISGKTDNLG	LDATVSMGIG	YTSDFGIGL	GYNFTYYNST	LGVRTPVLMV	NALNNNLRIA
130	140	150	160	170	180
IPIQIAASKD	PFGKYTISQY	KDYLGISTD	QIRYYTEIDV	FNQVRLYIKY	GQSGYKNVKN
190	200	210	220	230	240
NFDMFAQSFG	FETRLYFLNR	TIGNVNINPF	IKVSYNTALA	SSDVMVRAGE	SLVNTTYSKK
250	260	270	280	290	300
ENKWEKNPYN	VTAAAVLGLT	ANSDMLSLCV	EPSLGYNAVY	KGKYKTDISKY	YKVQHNLYWG
310	320	330	340	349	
AYAELYITPV	QDIEWYFEMD	INNGNSRQTS	SIPYFESTT	GITWYLPEF	

pTrep620--Copy 8

10	20	30	40	50	60
MTMITNRGSM	YGDQDDWIDF	LTDGNQLRAR	MDQLGFVLGN	NTIKGTFLGR	TQDAVTSLGS
70	80	90	100	110	120
IISGKTDNLG	LDATVSMGIG	YTSDFGIGL	GYNFTYYNST	LGVRTPVLMV	NALNNNLRIA
130	140	150	160	170	180
IPIQIAASKD	PFGKYTISQY	KDYLGISTD	QIRYYTEIDV	FNQVRLYIKY	GQSGYKNVKN
190	200	210	220	230	240
NFDMFAQSFG	FETRLYFLNR	TIGNVNINPF	IKVSYNTALA	SSDVMVRAGE	SLVNTTYSKK
250	260	270	280	290	300
ENKWEKNPYN	VTAAAVLGLT	ANSDMLSLCV	EPSLGYNAVY	KGKYKTDISKY	YKVQHNLYWG
310	320	330	340	350	360
AYAELYITPV	QDIEWYFEMD	INNGNSRQTS	SIPYFESTT	GITWYLPEF	

END
 pTrep620--Copy 8

-57-

Appendix #6--Predicted amino acid sequence of pTrep702
recombinant product

<-----Protein derived from
MetThrMetIleThrProSerLeuHisAlaCysArgSerThrLeuGluAspProArgVa
1 ATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGT
pUC polylinker----->||<---Copy 6 ORF--->
ProSerSerAsnTrpGluPheProMetAlaLeuLeuThrIleAlaSerAlaSerAlaPh
61 CCGAGCTCGAATTGGGAATTCCTATGGCATTATTGACTATAGCTAGTGCATCTGCTTT
GlyMetTyrGlyAspArgAspSerTrpIleAspPheLeuThrHisGlyAsnGlnPheAr
121 GGTATGTACGGCGACAGAGATTCTTGGATCGACTTCCTTACTCATGGTAATCAGTTTCAG
AlaArgMetAspGlnLeuGlyPheValLeuGlyAsnGlyThrIleLysGlyThrPheGl
181 GCTAGAATGGATCAATTAGGTTTCGTTTTCAGTAACGGTACTATTAAAGGTACTTTTCGG
PheArgSerGlnAlaIleGlyThrAlaLeuGlyAsnIleIleSerGlyAsnThrGlyAs
241 TTTAGATCTCAAGCTATTGGAACAGCATTAGGTAATATCATTTCAGGTAATACTGGAAA
ValAspLeuGlnThrThrIleSerAlaGlyIleGlyTyrThrSerGluProPheGlyIl
301 GTAGATTTACAACTACTATTTCTGCTGGTATAGGTTATACTTCTGAGCCTTTTCGGTAT
GlyValGlyTyrAsnTyrThrTyrValAsnProArgLeuGlyValHisThrProValLe
361 GCGGTAGGTTATACTACACTTATGTAAATCCTAGATTAGGCGTTCATACTCCTGTACT
MetIleAsnAlaLeuAsnAsnAsnLeuArgIleAlaValProValGlnIleAlaValSe
421 ATGATCAATGCTTTAAACAACAACCTTAAGAATAGCAGTTCCTGTTCAAATAGCTGTAAG
HisAspProPheAsnAspSerAlaLysPheProTyrSerSerSerThrLysAspTyrMe
481 CATGATCCTTTCAATGATTCTGCTAAATTCCTTATTCATCATCTACAAAAGATTATAT
GlyIleSerThrAspIleGlnLeuArgTyrTyrThrGlyIleAspAlaPheAsnAlaIl
541 GGTATAAGCACTGATATACAATTAAGATACTATACTGGTATAGATGCTTTCAATGCTAT
ArgLeuTyrPheLysTyrGlyGlnAlaGlyPheLysThrAlaAsnGlyAlaGlyAlaSe
601 AGATTATACTTCAAATACGGACAAGCTGGATTTAAACAGCTAACGGAGCTGGAGCTAG
GluTyrPheAlaGlnSerLeuGlyPheGluAlaArgPheTyrPheLeuAsnThrProVa
661 GAGTATTTTGCTCAGTCATTAGGTTTTGAAGCTAGATTCTATTTCTTGAATACTCCTGT
GlyAsnValThrIleAsnProPheIleLysValValTyrAsnThrAlaLeuLysGlyVa
721 GGAAACGTAACCTATCAATCCTTTCATCAAAGTTGTTTATAACACAGCTTTAAAGGTGT
SerArgThrValArgAlaGlyGluAlaValGlnAsnThrValSerGlyTyrHisProSe
781 AGCAGAACTGTAAGAGCTGGAGAAGCTGTACAAAATACTGTTTCTGGTTATCATCCTTC
AsnProAsnTyrLysLeuAspAlaPheAlaGlyArgTyrIleGlyLysAspPheLysTr
841 AATCCTAATTATAAATTAGATGCATTTGCTGGTAGATACATTGGTAAAGATTTCAAATG
AspSerAsnProTyrAspValLysAlaGlnAlaValLeuGlyIleThrAlaAsnSerAs
901 GATTCAAATCCTTATGATGTAAAAGCTCAGGCTGTATTAGGTATCACTGCTAACAGCGA
ValValSerLeuTyrValGluProSerLeuGlyTyrGlnAlaThrTyrLeuGlyLysAs
961 GTAGTATCTCTTTATGTTGAGCCTTCTTTAGGTTATCAAGCTACATATTTAGGAAAAAA

-68-

---End of Copy 6 ORF --->||<---beta

IleSerGluAsnProTyrLeuAsnIleAspSerLysValGlnHisArgGluPheGlyAsn
1021 ATATCTGAAAATCCATATTTAAATATAGATTCTAAAGTACAACATAGGGAATTCGGCAAT

galactosidase--->

SerLeuAlaValValLeuGlnArgArgAspTrpGluAsnProGlyValThrGlnLeuAsn
1081 TCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACCTTAAT

ArgLeuAlaAlaHisProProPheAlaSerTrpArgAsnSerGluGluAlaArgThrAsp
1141 CGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGAT

ArgProSerGlnGlnLeuArgSerLeuAsnGlyGluTrpArgLeuMetArgTyrPheLeu
1201 CGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCCTGATGCGGTATTTCTC

LeuThrHisLeuCysGlyIleSerHisArgIleTrpCysThrLeuSerThrIleCysSer
1261 CTTACGCATCTGTGCGGTATTTACACCGCATATGGTGCACTCTCAGTACAATCTGCTCT

AspAlaAlaAM

1321 GATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGG

1381 GCTTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATG

-69-

Appendix #7--Predicted protein sequence of pTrep704 recombinant product

<-----Protein derived from

MetThrMetIleThrProSerLeuHisAlaCysArgSerThrLeuGluAspProArgVal

1 ATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTA

pUC polylinker----->||<-----Copy 6 ORF

ProSerSerAsnTrpGluPheProMetAlaLeuLeuThrIleAlaSerAlaSerAlaPhe

61 CCGAGCTCGAATTGGGAATTCCTATGGCATTATTGACTATAGCTAGTGCATCTGCTTTC

GlyMetTyrGlyAspArgAspSerTrpIleAspPheLeuThrHisGlyAsnGlnPheArg

121 GGTATGTACGGCGACAGAGATTCTTGGATCGACTTCCTTACTCATGGTAATCAGTTCAGA

AlaArgMetAspGlnLeuGlyPheValLeuGlyAsnGlyThrIleLysGlyThrPheGly

181 GCTAGAATGGATCAATTAGGTTTCGTTTTAGGTAACGGTACTATTAAAGGTACTTTCCGT

PheArgSerGlnAlaIleGlyThrAlaLeuGlyAsnIleIleSerGlyAsnThrGlyAsn

241 TTTAGATCTCAAGCTATTGGAACAGCATTAGGTAATATCATTTTCAGGTAATACTGGAAAT

ValAspLeuGlnThrThrIleSerAlaGlyIleGlyTyrThrSerGluProPheGlyIle

301 GTAGATTTACAACTACTATTTCTGCTGGTATAGGTTATACTTCTGAGCCTTTCGGTATT

GlyValGlyTyrAsnTyrThrTyrValAsnProArgLeuGlyValHisThrProValLeu

361 GGCGTAGGTTATACTACACTTATGTAAATCCTAGATTAGGCGTTCATACTCCTGTACTT

MetIleAsnAlaLeuAsnAsnAsnLeuArgIleAlaValProValGlnIleAlaValSer

421 ATGATCAATGCTTTAACAACAACCTTAAGAATAGCAGTTCCTGTTCAAATAGCTGTAAGT

HisAspProPheAsnAspSerAlaLysPheProTyrSerSerSerThrLysAspTyrMet

481 CATGATCCTTTCAATGATTCTGTAAATTCCTTATTCATCATCTACAAAAGATTATATG

GlyIleSerThrAspIleGlnLeuArgTyrTyrThrGlyIleAspAlaPheAsnAlaIle

541 GGTATAAGCACTGATATACAATTAAGATACTATACTGGTATAGATGCTTTCAATGCTATA

ArgLeuTyrPheLysTyrGlyGlnAlaGlyPheLysThrAlaAsnGlyAlaGlyAlaSer

601 AGATTATACTTCAAATACGGACAAGCTGGATTTAAACAGCTAACGGAGCTGGAGCTAGT

GluTyrPheAlaGlnSerLeuGlyPheGluAlaArgPheTyrPheLeuAsnThrProVal

661 GAGTATTTTGCTCAGTCATTAGGTTTTGAAGCTAGATTCTATTTCTTGAATACTCCTGTT

GlyAsnValThrIleAsnProPheIleLysValValTyrAsnThrAlaLeuLysGlyVal

721 GGAAACGTAACCTATCAATCCTTTTCATCAAAGTTGTTTATAACACAGCTTTAAAGGTGTA

SerArgThrValArgAlaGlyGluAlaValGlnAsnThrValSerGlyTyrHisProSer

781 AGCAGAACTGTAAGAGCTGGAGAAGCTGTACAAAATACTGTTTCTGGTTATCATCCTTCT

AsnProAsnTyrLysLeuAspAlaPheAlaGlyArgTyrIleGlyLysAspPheLysTrp

841 AATCCTAATTATAAATTAGATGCATTTGCTGGTAGATACATTGGTAAAGATTTCAAATGG

AspSerAsnProTyrAspValLysAlaGlnAlaValLeuGlyIleThrAlaAsnSerAsp

901 GATTCAAATCCTTATGATGTAAAGCTCAGGCTGTATTAGGTATCACTGCTAACAGCGAT

ValValSerLeuTyrValGluProSerLeuGlyTyrGlnAlaThrTyrLeuGlyLysAsn

961 GTAGTATCTCTTTATGTTGAGCCTTCTTTAGGTTATCAAGCTACATATTTAGGAAAAAAC

-70-

IleSerGluAsnProTyrLeuAsnIleAspSerLysValGlnHisSerLeuAlaTrpGly
1021 ATATCTGAAAATCCATATTTAAATATAGATTCTAAAGTACAACATAGCTTAGCTTGGGGT
AlaTyrAlaGluLeuTyrValArgProValGlnAspLeuGluTrpTyrPheGluMetAsp
1081 GCTTATGCAGAACTTTATGTAAGACCTGTTCAAGATCTTGAATGGTACTTCGAGATGGAT
ValAsnAsnGlyGlyThrArgGlnGluSerGlyIleProValTyrPheLysSerThrThr
1141 GTTAATAATGGCGGTACAAGACAAGAATCTGGTATCCCTGTATACTTTAAATCTACTACA
GlyIleThrTrpTyrLeuProAlaPheAsnOC
1201 GGTATAACTTGGTATTTACCTGCTTTCAATTAATTAGAAGTTAATTAATAGAAATTAATG

-71-

Appendix #8--Predicted protein sequence of recombinant product
of pTrep505

<---Protein derived from

MetThrMetIleThrProSerLeuHisAlaCysArgSerThrLeuGluAspProArgVal
 1 ATGACCATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTA

pUC polylinker----->||<---Copy 2 ORF --->
 ProSerSerAsnTrpGluPheProArgMetAspGlnLeuGlyPheValLeuGlyAsnSer
 61 CCGAGCTCGAATTGGGAATTCCTAGAAATGGATCAATTAGGATTGTTTTAGGTAATAGC

ThrIleLysGlyThrPheGlyPheArgSerGlnSerLeuSerThrGlnLeuGlyTyrIle
 121 ACCATTAAAGGTACTTTTCGGTTTTAGATCTCAGAGTTTATCAACTCAATTAGGATATATT

LeuAlaIleTyrLysAspTyrThrTyrLeuGlyAlaThrIleSerGlyGlyIleGlyTyr
 181 TTGGCTATATATAAAGATTATACTTATTTAGGAGCAACTATTTCCGGCGGTATAGGATAT

ThrSerGluAlaPheSerIleGlyLeuGlyTyrAsnTyrThrThrProLeuProIleSer
 241 ACTTCTGAGGCTTTTAGTATAGGTTTAGGTTATAATTATACTACACCGCTTCCTATTAGT

TyrAsnPheGlySerHisThrProValLeuMetIleAsnAlaLeuAsnAspAsnLeuArg
 301 TATAACTTTGGTTCTCATACTCCTGTACTTATGATTAATGCTTTAAATGATAATTTGAGG

IleValIleProValGlnIleLeuValHisAspGlyAsnMetAsnMetThrAspAsnIle
 361 ATAGTTATTCCTGTACAAATATTAGTACATGATGGTAATATGAATATGACGGATAATATT

AsnTyrLeuTyrAsnPheLeuGlyIleSerThrAspThrGlnIleArgTyrTyrThrGly
 421 AATTATTTATATAATTTTTTAGGTATAAGTACTGATACTCAAATAAGATATTATACAGGC

IleAspAlaPheAsnGluIleArgLeuTyrValLysTyrGlyGlnLeuGlyTyrLysGly
 481 ATAGACGCTTTTAATGAAATAAGATTATATGTAAATACGGACAATTAGGATATAAAGGC

GlySerTyrThrAspLysSerTyrAspGluGluPhePheAlaArgSerPheGlyPheGlu
 541 GGTTCATATACGGATAAAAGTTATGATGAAGAATTTTTTGCAAGATCATTTGGTTTTGAA

ThrArgPheTyrPheLeuAsnThrAlaValGlyAsnValThrIleAsnProPheIleLys
 601 ACAAGATTCTATTTTTTGAATACTGCTGTTGGAAATGTAATATCAATCCTTTTATTAAA

ValAlaTyrAsnThrAlaLeuHisGlyPheSerThrMetValArgSerLeuAspSerVal
 661 GTAGCATATAATACAGCTTTGCATGGATTAGTACTATGGTAAGATCATTAGATAGTGTC

IleGluGluIleGluGlyTyrSerSerAspArgThrAlaLysAlaAlaGlyAsnIleAsn
 721 ATTGAAGAAATAGAGGTTATAGTTCAGATCGTACCGCTAAAGCAGCAGGAAATATTAAT

AlaLysTrpAspLysAsnProTyrAspValThrValGlnAlaValLeuGlyValThrAla
 781 GCTAAATGGGATAAGAATCCTTATGATGTAAGTGTGCAGGCAGTATTGGGAGTAACTGCT

AsnSerAspIleValSerLeuTyrValGluProSerLeuGlyTyrArgAlaLysTyrLeu
 841 AATAGCGATATAGTATCACTTTATGTTGAGCCTTCTTTAGGTTATAGGGCTAAATATTTA

GlyLysLeuThrTyrGluAspProAspGlyLysValAsnPheAspPheLysValAsnHis
 901 GGAAAATTAACATATGAAGATCCAGATGGAAAAGTTAATTTTTGATTTTAAAGTTAATCAT

TyrLeuSerTrpCysAlaTyrAlaGluLeuTyrIleThrProValLysAspLeuGluTrp
 961 TATTTATCTTGGTGTGCTTATGCAGAGCTTTATATAACACCTGTAAAAGATTTAGAATGG

TyrPheGluMetAspValAsnAsnSerAspSerAspSerThrGlyIleProValSerPhe
 1021 TATTTTGAATGGATGTTAATAAGTATGATTGAGATTCTACAGGTATACCTGTTAGTTTT

-72-

AlaSerThrThrGlyIleThrTrpTyrLeuProGluPheOC OC
1081 GCTTCTACTACAGGAATAACTTGGTATTTGCCAGAATTTTAATTATAAAGCAAATTTTAT
1141 ATGACAAAATAAAAAATGGGGCATTATTATTAAAAAATAAATACCCACATTTTATTAA
1201 ATAACCTCTTAAATAATTTTACA4TTTATATTTTATTAGTATAATAAAATATAAAGTTAA
1261 ATTTAGGTGTGTACAATGAAAAAAGTTTCTAATTATGACAGTATTATTAAGTATGTCA
End of
1321 TATTGTTCAATATTTGGTATGTATGGACATCAGGACGATTGGATTGATTTTCTTACAGTC
T. hyo. insert----->|
1381 GGTAATCAGTTTAGAGGGAATTC

What is Claimed is:

1. A protein capable of eliciting at least one antibody capable of recognizing at least one epitope of at least one T. hyo. antigen having a molecular weight of about 39 kDa.

2. The protein of Claim 1 encoded by gene 1.

3. The protein of Claim 2 encoded by the full length gene 1.

4. The protein of Claim 1 encoded by gene 2.

5. The protein of Claim 1 encoded by gene 3.

6. The protein of Claim 1 encoded by gene 4.

7. The protein of Claim 1 encoded by gene 5.

8. The protein of Claim 1 encoded by gene 6.

9. The protein of Claim 1 encoded by gene 7.

10. The protein of Claim 1 encoded by gene 8.

11. The protein of Claim 1 encoded by the full length gene 2.

12. The protein of Claim 1 encoded by the full length gene 3.

13. The protein of Claim 1 encoded by the full length gene 4.

14. The protein of Claim 1 encoded by the full length gene 5.

15. The protein of Claim 1 encoded by the full length gene 6.

16. The protein of Claim 1 encoded by the full length gene 7.

17. The protein of Claim 1 encoded by the full length gene 8.

18. DNA encoding at least one protein capable of eliciting at least one antibody recognizing at least one epitope of at least one T. hyo. antigen having a molecular weight of about 39 kDa.

19. The DNA of Claim 16 wherein said DNA is selected from the group consisting of genes 1, 2, 3,

-74-

4, 5, 6, 7 and 8 encoding a T. hyo. 39 kDa protein and mixtures thereof.

20. The DNA sequence of Claim 18 wherein the gene is a full length gene.

21. A host genetically engineered with DNA of Claim 18.

22. A host genetically engineered with DNA of Claim 19.

23. A host genetically engineered with DNA of Claim 20.

24. An expression vehicle including DNA of Claim 18.

25. An expression vehicle including the DNA of Claim 19.

26. An expression vehicle including the DNA of Claim 20.

27. Protein expressed by the host of Claim 21.

28. Protein expressed by the host of Claim 22.

29. Protein expressed by the host of Claim 23.

30. A process for protecting an animal against T. hyo., comprising:

administering to an animal to be protected an effective amount of at least one protein of Claim 1.

31. A process for protecting an animal against T. hyo., comprising:

administering to an animal to be protected an effective amount of at least one protein of Claim 27.

32. A vaccine for protecting an animal against T. hyo. comprising:

at least one protein of Claim 1 in conjunction with a pharmaceutically acceptable carrier.

33. A vaccine for protecting an animal against T. hyo. comprising:

at least one protein of Claim 27 in conjunction with a pharmaceutically acceptable carrier.

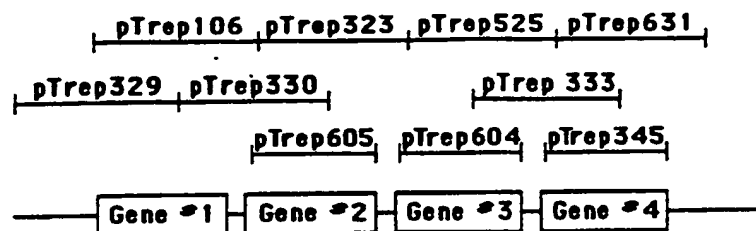
-75-

34. Antibody which recognizes T. hyo antigen
having a molecular weight of about 39 kDa.

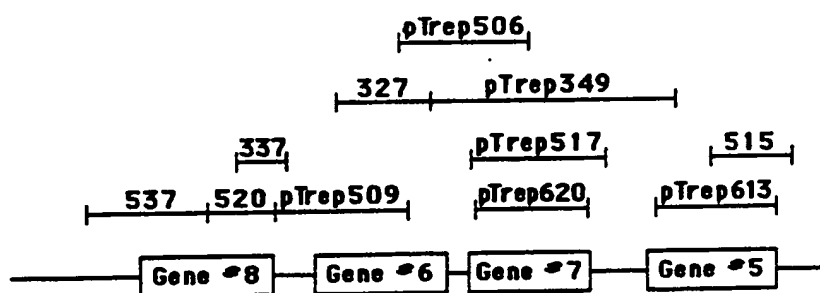
1/5

Figure #1--Cloning and sequencing of 39 kDa genes 1-8

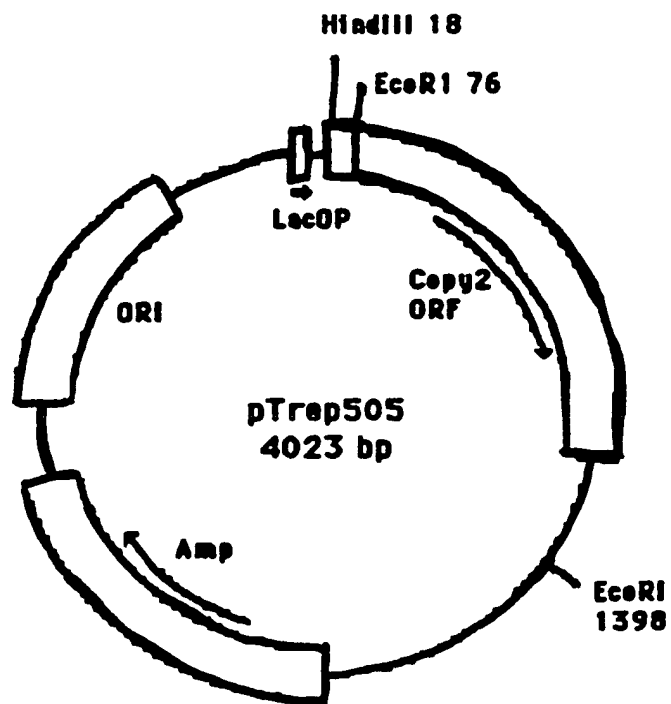
GENES 1-4



GENES 5-8

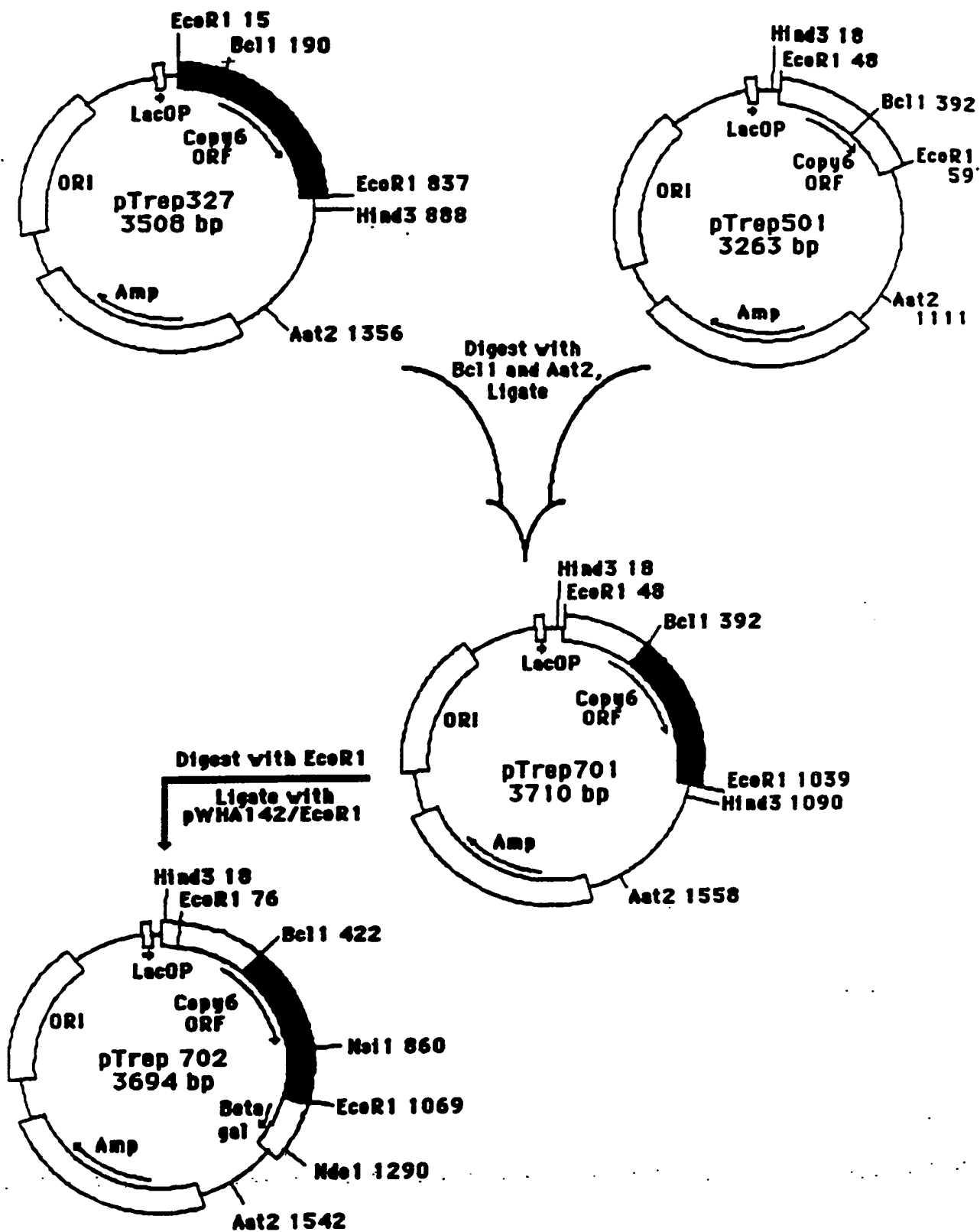
1 Kb

2/5

Figure #2--Schematic of pTrep505

3/5

Figure #3--Construction of pTrep702



4/5

Figure #4--Construction of pTrep704

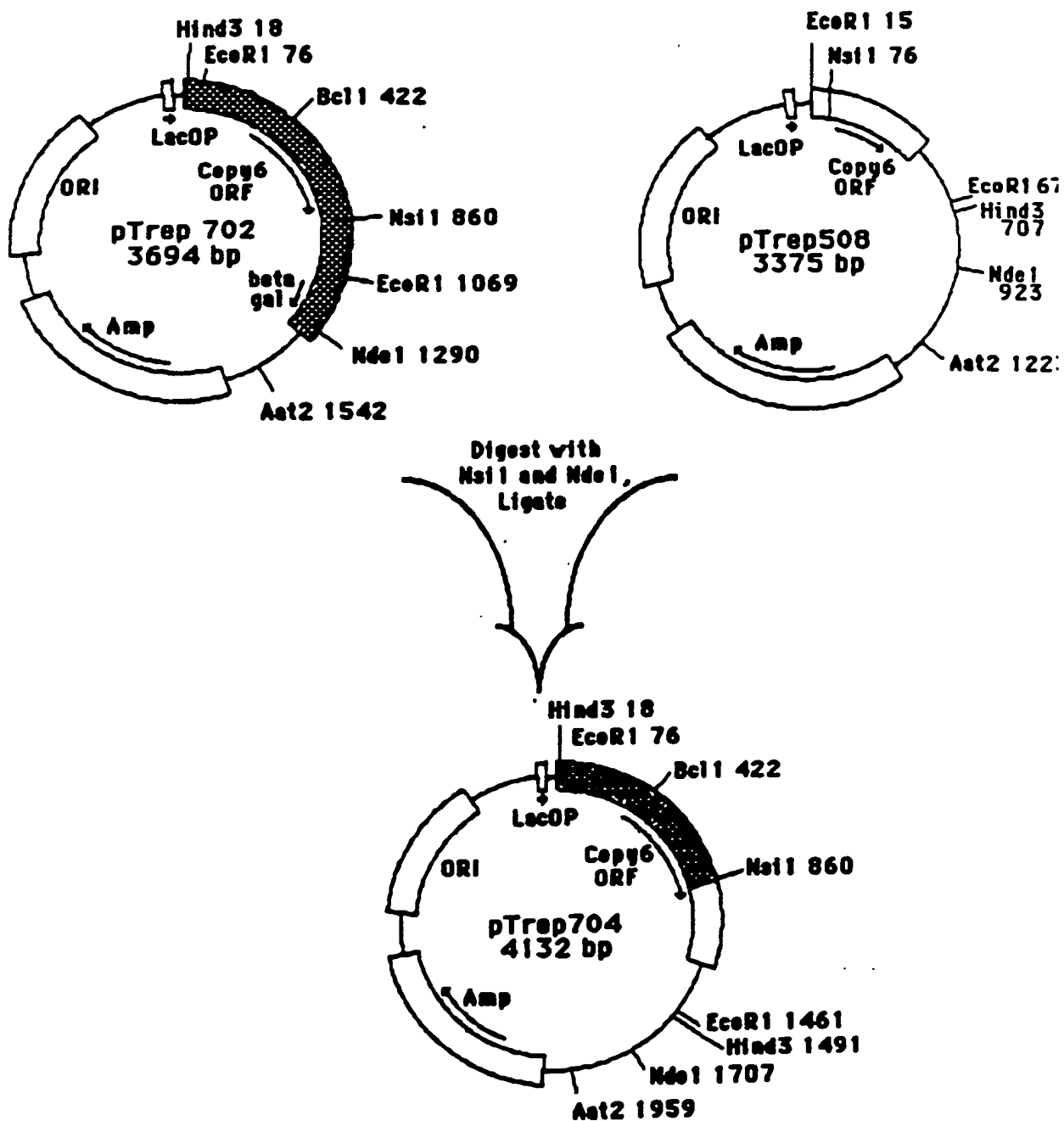
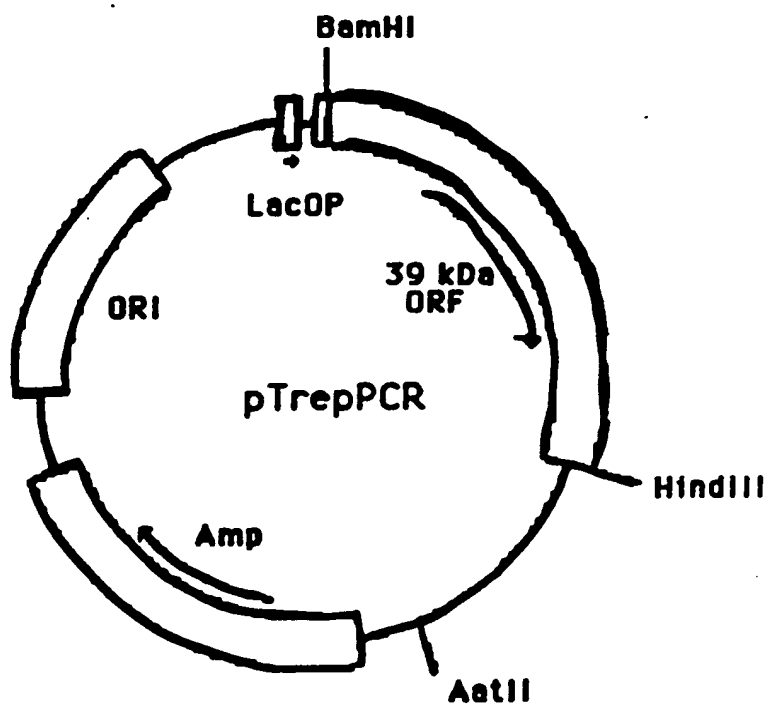


Figure #5--Schematic of PCR derived ^{5/5} expression clones



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.